



Etude structurale d'un complexe de trois protéines de la division du pneumocoque, DivIB, DivIC et FtsL

Soizic Masson

► To cite this version:

Soizic Masson. Etude structurale d'un complexe de trois protéines de la division du pneumocoque, DivIB, DivIC et FtsL. Biophysique [physics.bio-ph]. Université Joseph-Fourier - Grenoble I, 2008. Français. NNT: . tel-00352348

HAL Id: tel-00352348

<https://theses.hal.science/tel-00352348>

Submitted on 12 Jan 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

*Thèse pour l'obtention du grade de
Docteur de l'Université Joseph Fourier*

Spécialité : Physique pour les Sciences du Vivant

Etude structurale d'un complexe de trois
protéines de la division du pneumocoque,
DivIB, DivIC et FtsL

Structural study of a complex of three proteins of the
Pneumococcus division, DivIB, DivIC and FtsL

Soutenue le 14 novembre 2008

Devant le jury composé de :

M. Franz Bruckert
M. Philippe Dumas
M. François-Xavier Barre
Mme Stéphanie Finet
M. Frank Gabel
M. Thierry Vernet

Président du jury
Rapporteur
Rapporteur
Examineur
Co-directeur de thèse
Examineur

Thèse préparée au Laboratoire d'Ingénierie des Macromolécules de l'Institut de
Biologie Structurale, sous la direction de André Zapun

« *Ce qui ne tue pas, rend plus fort.* »
Friedrich Nietzsche

A Cécile,

Remerciements

Tout d'abord, merci à ceux, grâce à qui j'ai eu cette brillante idée de commencer une thèse : Benoît Guillon et Nicolas Thiellens, Michel Masson, Franz Bruckert, et bien sûr, André Zapun et Thierry Vernet.

Mais surtout, merci à ceux, grâce à qui j'ai réussi à la finir !

Merci notamment à la cellule de soutien psychologique qui s'est mise en place durant ces trois années et demie: Cécile Frolet, Anne Marie Di Guilmi, Lamya El Mortaji, Florent Bernaudat, Benoit Gallet.

Merci à ceux qui en font partie depuis maintenant fort longtemps : Claire Magimel, Muriel Moutard, Elsa Mas, Yan Bayssat, Maryse Masson.

Je tiens aussi à remercier toutes les personnes avec qui j'ai travaillé dans le cadre de cette thèse. Cela a vraiment été un plaisir de travailler avec vous tous, et bien sûr, je vous remercie aussi pour avoir bien voulu travailler avec moi !

Je voudrais tout particulièrement remercier Frank Gabel pour le temps, l'énergie que tu as consacré à ce projet, mais aussi et surtout ta persévérance dans les moments difficiles et ton amitié.

Thierry, merci pour la confiance que vous avez eue en moi et en ce projet : bien souvent vous avez plus cru en moi, que moi-même !

André, merci pour toutes ces discussions, pour tes points de vue si radicalement différents des miens !

Merci à vous deux pour votre écoute, votre respect de mes idées et envies. Honnêtement, vous êtes tous les deux des sacrés personnages, et je suis vraiment très heureuse de vous avoir rencontrés, professionnellement et humainement parlant.

A tout le LIM, les permanents, ceux qui sont passés, ceux qui viennent d'arriver, merci pour les conseils, l'aide, le soutien, mais surtout pour cette ambiance extraordinaire. Pas une seule journée, même dans mes grands moments de doutes (et il y en a eu...), je n'ai traîné des pieds pour venir au labo. Cet enthousiasme, je vous le dois.

Je tiens aussi à dire à toutes les personnes que j'ai eu l'occasion de rencontrer à l'Institut durant cette thèse, qu'elles ont largement participé à rendre ces trois années passionnantes, et je les remercie pour ça.

Je garde de vous tous un souvenir très précieux.

Avertissement au lecteur

Les travaux présentés dans ce manuscrit ont été réalisés par moi-même au sein du Laboratoire d'Ingénierie des Macromolécules de l'IBS, à l'exception des travaux sous cités.

Les expériences de RMN ont été réalisées en collaboration avec Jean-Pierre Simorre (LRMN/IBS) et Thomas Kern (LRMN/IBS) au Laboratoire de Résonance Magnétique Nucléaire de l'IBS.

Les expériences de diffusion dynamique de lumière ont été réalisées au Laboratoire des Protéines Membranaires de l'IBS.

Les expériences d'ultracentrifugation analytique ont été réalisées avec Christine Ebel (LBM/IBS) au Laboratoire de Biophysique Moléculaire de l'IBS.

Les marquages isotopiques au deutérium des protéines EC et KL ont été réalisés au Laboratoire de Deutération de l'ILL, avec Martine Moulin (D-Lab/ILL) et Michael Haertlein (D-Lab/ILL).

Les expériences de diffusion aux petits angles de neutrons ont été réalisées à l'ILL sur l'instrument D22, en collaboration avec Frank Gabel (LBM/IBS) et Phil Callow (ILL).

Les expériences de diffusion aux petits angles de rayons X ont été réalisées à l'ESRF, sur l'instrument ID02, en collaboration avec Frank Gabel (LBM/IBS), avec l'aide de Anuj Shukla (ESRF), Theyencheri Narayanan (ESRF) et Stéphanie Finet (UPMC).

Les expériences d'analyse d'interaction de biomolécules par résonance plasmonique de surface ont été réalisées avec l'assistance de Nicole Thielens (LEM/IBS) au Laboratoire d'Enzymologie Moléculaire.

Les analyses par spectrométrie de masse ont été réalisées au Laboratoire de Spectrométrie de Masse des Protéines de l'IBS, par Bernard Dublet (LSMP/IBS).

Les analyses en acides aminés et les séquençages N-terminaux ont été réalisés au laboratoire d'Enzymologie Moléculaire, par Jean-Pierre Andrieu (LEM/IBS).

Institut de Biologie Structurale Jean-Pierre Ebel
UMR 5075
41, rue Jules Horowitz
38027 Grenoble Cedex 1

Sommaire

Remerciements	p.5
Avertissement au lecteur	p.7
Listes des figures et tableaux	p.13
Liste des abréviations utilisées	p.17

Introduction	p.19
A. Contexte de l'étude	p.19
I. Des bactéries et des hommes	p.19
II. Un exemple de bactérie: le pneumocoque	p.21
III. La paroi bactérienne	p.23
IV. La division bactérienne	p.25
IV.1. La constriction de la membrane	p.27
IV.2. Le transport des chromosomes	p.29
IV.3. La synthèse de la paroi septale	p.31
V. Trois protéines de la division : FtsL, DivIC et DivIB	p.33
V.1. FtsL	p.33
V.1.1. Quelques données génétiques	p.33
V.1.2. Analyse de la séquence primaire	p.33
V.1.3. Quelques données biochimiques	p.34
V.2. DivIC	p.35
V.3. Un complexe entre FtsL et DivIC	p.36
V.4. DivIB	p.37
V.4.1. Quelques données génétiques	p.37
V.4.2. Analyse de la séquence primaire	p.38
V.4.3. Analyse structurale	p.39
V.4.3.1. Chez <i>G. stearothermophilus</i> , bactérie Gram-positif	p.39
V.4.3.2. Chez <i>E. coli</i> , bactérie Gram-négatif	p.40
V.4.4. Quelques données biochimiques et fonctionnelles	p.42
V.5. Un complexe entre FtsL, DivIC et DivIB	p.43
VI. Synthèse	p.44
B. Les outils de la physique	p.45
I. Les microscopies	p.45
II. La diffusion/diffraction	p.47
II.1. la cristallographie	p.48
II.2. La diffusion aux petits angles	p.49
II.2.1. La formule de Debye	p.49
II.2.2. La fonction de distribution de distances	p.52
II.2.3. La complémentarité rayons X/neutrons	p.53
III. Les spectroscopies	p.55
III.1. La spectroscopie optique	p.55
III.2. La RMN	p.56
III.2.1. Les fréquences de résonance	p.56
III.2.2. Les spectres multidimensionnelle	p.58
III.2.3. La relaxation	p.59
IV. Le comportement hydrodynamique	p.61
IV.1. La diffusion dynamique de lumière (DLS)	p.61
IV.2. L'ultracentrifugation analytique	p.62
V. L'analyse d'interaction de biomolécules par résonance plasmonique de surface	p.64
C. Présentation du modèle et objectif du travail	p.67
I. Présentation du modèle	p.67
II. Objectif du travail	p.69

Chapitre 1 : Le complexe contraint EC/KL	p.71
I. Co-purification du complexe EC/KL	p.73
II. Analyse structurale et biochimique	p.75
II.1. Première approche par RMN	p.75
II.2. Insertion d'un résidu aromatique dans la protéine KL	p.75
II.3. Estimation de la taille du complexe EC/KL (DLS, SEC, AUC)	p.77
III. Elaboration des constructions génétiques et des protocoles de purification	p.81
III.1. marquages isotopiques	p.81
III.2. Co-expression	p.83
IV. Etude par SANS	p.85
IV.1. Caractérisation des échantillons par AUC et DLS	p.85
IV.2. Détermination de l'état oligomérique du complexe EC/KL	p.85
IV.3. Détermination de la forme du complexe EC/KL	p.87
IV.4. Détermination de l'arrangement des protéines EC et KL dans le complexe EC/KL	p.89
V. Etude par SAXS	p.91

Chapitre 2. Le domaine extracellulaire de DivIB de <i>S. pneumoniae</i>	p.95
I. Analyse préliminaire de la partie extracellulaire de DivIB (RMN et protéolyse limitée)	p.97
I.1. Analyse de la structuration	p.97
I.2. Délimitation en trois domaines	p.99
II. Etude structurale des domaines de la partie extracellulaire de DivIB	p.101
II.1. Production des variants tronqués de la partie extracellulaire de DivIB	p.101
II.2. Caractérisation de l'auto-association des protéines	p.103
II.3. Etude structurale du domaine β par RMN	p.105
II.4. Etude structurale par SAXS	p.107
II.4.1. préparation des échantillons	p.107
II.4.2. Etude du domaine β	p.107
II.4.3. Etude de la protéine $\beta\gamma$	p.111
II.4.4. Etude de la protéine $\alpha\beta\gamma$ (ou IB)	p.115

Chapitre 3 : l'interaction entre le complexe EC/KL et le domaine extracellulaire de DivIB	p.117
I. Etude sur la partie extracellulaire de DivIB complexée avec le complexe EC/KL, par RMN	p.119
I.1. Analyse préliminaire	p.119
I.2. Détermination de l'épitope du domaine β impliqué dans l'interaction	p.121
II. Cartographie des domaines d'interaction du domaine extracellulaire de DivIB et du complexe EC/KL par analyse d'interaction de biomolécule par résonance plasmonique de surface	p.123
II.1. Mesure de l'interaction entre les domaines de la partie extracellulaire de DivIB et le complexe EC/KL	p.123
II.2. Cartographie des protéines EC et KL	p.125
III. Modélisation du complexe trimérique par SANS	p.129
III.1. première analyse des données	p.129
III.2. Présentation d'un modèle d'association entre $\beta\gamma$ et EC/KL	p.131
III.3. Exploitation de la courbe de diffusion	p.133
III.3.1. Modélisation <i>ab initio</i>	p.133
III.3.2. Ajustement des corps rigides	p.137

Chapitre 4 : Matériel et méthodes	p.141
I. Constructions génétiques	p.141
I.1. Le complexe contraint EC/KL	p.141
I.2. La partie extracellulaire de DivIB	p.142
II. Production des protéines	p.143
II.1. Production classique	p.143

II.2. Production pour marquage isotopique	p.144
II.2.1. Marquage à l'isotope 15 de l'azote	p.144
II.2.2. Marquage à l'isotope 2 de l'hydrogène, le deutérium D	p.145
II.3. Production en bioréacteur	p.146
III. Purification	p.146
III.1. Purification du complexe contraint EC/KL et ses variants	p.146
III.1.1 Purification de (Strep)EC	p.146
III.1.2 Purification de (his6)EC, KL, EC* et KL*	p.146
III.1.3. Préparation de EC/KL, EC*/KL et EC/KL*	p.147
III.1.4. Co-purification de EC/KL	p.147
III.2. Purification du domaine extracellulaire de DivIB et ses variants	p.148
III.2.1 Purification de $\alpha\beta\gamma$ (ou IB), $\beta\gamma$, β , par chromatographie d'affinité	p.148
III.2.2. Préparation d'échantillon de protéine β par protéolyse limitée de protéine $\alpha\beta\gamma$ (ou IB)	p.148
III.2.3. Tests de purification des protéines α , β , $\alpha\beta$ et $\beta\gamma$	p.149
III.2.4. Purification par chromatographie échangeuse d'ions	p.149
III.2.5. Purification par chromatographie d'exclusion de taille	p.150
III.3. Immuno-révélation	p.151
IV. Expériences	p.151
IV.1. RMN	p.151
IV.2. Diffusion dynamique de lumière (DLS)	p.152
IV.3. Chromatographie par exclusion de taille	p.153
IV.4. Ultracentrifugation analytique	p.154
IV.5.SANS et SAXS	p.155
IV.5.1. Enregistrement des courbes de diffusion	p.155
IV.5.2. Extraction des valeurs d'intensité à l'origine et du rayon de giration	p.156
IV.5.3. Calibration de l'échelle d'intensité	p.156
IV.5.4. Fonction de distribution de paires	p.157
IV.5.5. Calcul des simili-courbes de diffusion aux petits angles	p.157
IV.5.6. Modélisation <i>ab initio</i>	p.157
IV.5.7. Modélisation avec des corps rigides (RBM)	p.157
IV.6. Protéolyse limitée	p.158
IV.7. Biacore	p.158
IV.7.1.Acquisition des sensorgrammes	p.158
IV.7.2. Traitement des sensorgrammes	p.159
Discussion	p.161
I. Le complexe contraint EC/KL	p.161
II. La partie extracellulaire de DivIB	p.164
III. Etude de l'interaction entre le complexe EC/KL et la partie extracellulaire de DivIB	p.167
III.1. Du côté du complexe EC/KL	p.167
III.2. Du côté de la partie extracellulaire de DivIB	p.168
III.3. A propos de l'interaction entre EC/KL et la partie extracellulaire de DivIB	p.170
Conclusion et perspectives	p.173
Bibliographie	p.177
Annexe	p.187
Résumé	p.190

Listes des figures

Figure 1 : Image de <i>S. pneumoniae</i> , par microscopie électronique.....	p.22
Figure 2 : La paroi bactérienne des bactéries à Gram négatif et à Gram positif.....	p.23
Figure 3 : Ordres de recrutement de protéines impliquées dans la division, au site de division chez <i>E. coli</i> et <i>B. subtilis</i>	p.26
Figure 4 : Assemblage des protéines FtsZ et FtsA à la membrane plasmique.....	p.28
Figure 5 : Le complexe hexamerique de FtsK à la membrane plasmique.....	p.30
Figure 6 : Assemblage de FtsW et FtsI à la membrane plasmique.....	p.32
Figure 7 : Alignement des séquences primaires de FtsL de <i>S. pneumoniae</i> , <i>B. subtilis</i> et <i>E. coli</i>	p.34
Figure 8 : Alignement des séquences primaires de DivIC de <i>S. pneumoniae</i> , <i>B. subtilis</i> et <i>E. coli</i>	p.36
Figure 9 : Alignement des séquences primaires de DivIB de <i>S. pneumoniae</i> , <i>B. subtilis</i> et <i>E. coli</i>	p.38
Figure 10 : Structure tridimensionnelle d'une partie de la partie extracellulaire de DivIB de <i>G. stearothermophilus</i>	p.40
Figure 11 : Alignement des séquences primaires des domaines structurés de DivIB de <i>G. stearothermophilus</i> et de <i>E. coli</i>	p.41
Figure 12 : Structure tridimensionnelle d'une partie du domaine extracellulaire de DivIB de <i>E. coli</i> et comparaison avec la structure d'une partie du domaine extracellulaire de <i>G. stearothermophilus</i>	p.41
Figure 13 : Assemblage de protéines de la division bactérienne.....	p.44
Figure 14 : Quelques ordres de grandeur de la biologie.....	p.45
Figure 15 : Représentation schématique de la diffusion d'un faisceau incident I_0 par un objet.....	p.49
Figure 16 : Représentation schématique de deux pics de résonance de RMN, de largeur différente...	p.60
Figure 17 : Principe de l'analyse d'interaction de biomolécules par résonance plasmonique de surface.....	p.66
Figure 18 : Représentation schématique du « coiled-coil » formé par le motif e et le motif k.....	p.67
Figure 19 : Construction des protéines EC et KL.....	p.68
Figure 20 : Schéma représentant le système des protéines utilisées.....	p.69
Figure 21 : Schéma des protéines élaborées à partir des domaines extracellulaires de DivIC et FtsL de <i>S. pneumoniae</i>	p.72
Figure 22 : Co-purification du complexe EC/KL.....	p.72
Figure 23 : Spectre HSQC [^{15}N ; ^1H] de la protéine KL dans le complexe EC/KL.....	p.74

Figure 24 : Ultracentrifugation analytique, vitesses de sédimentation d'échantillons de complexe EC/KL.....	p.76
Figure 25 : Estimation de la pureté des échantillons de complexe EC/KL dans les différentes combinaisons de marquage isotopique.....	p.84
Figure 26 : Ultracentrifugation analytique d'échantillons de complexe EC/KL hydrogéné et partiellement deutérés.....	p.84
Figure 27 : Comparaison entre la fonction de distributions de paires normalisée du complexe EC/KL et celles d'un cylindre et d'une sphère	p.88
Figure 28 : Comparaison des courbes de diffusion d'échantillons de complexe EC/KL, hydrogéné en tampon 100% H ₂ O et partiellement deutérés en tampon 41% D ₂ O/59% H ₂ O	p.90
Figure 29 : Comparaison des fonctions de distributions de paires normalisées des complexes EC/KL hydrogéné et partiellement deutérés.....	p.90
Figure 30 : Evolution de la masse apparente du complexe EC/KL en fonction de la concentration.....	p.92
Figure 31 : Spectre HSQC [¹⁵ N ; ¹ H] de la partie extracellulaire de DivIB de <i>S. pneumoniae</i>	p.96
Figure 32 : Protéolyse limitée de la partie extracellulaire de DivIB de <i>S. pneumoniae</i>	p.96
Figure 33 : Séquence primaire de DivIB de <i>S. pneumoniae</i>	p.98
Figure 34 : Spectre HSQC [¹⁵ N ; ¹ H] du domaine β de DivIB de <i>S. pneumoniae</i>	p.98
Figure 35 : Schéma des protéines variantes de DivIB de <i>S. pneumoniae</i>	p.100
Figure 36 : Bilan des purifications des différentes constructions de la partie extracellulaire de DivIB de <i>S. pneumoniae</i>	p.102
Figure 37 : Contrôle de la qualité des échantillons utilisés pour les expériences de SAXS.....	p.106
Figure 38 : Comparaison des courbes de diffusion aux petits angles du domaine β de différentes espèces de bactéries.....	p.108
Figure 39 : Comparaison des courbes de diffusion des domaines βγ, et β de <i>S. pneumoniae</i> , avec la simili-courbe de diffusion du domaine β d' <i>E. coli</i>	p.110
Figure 40 : Modèle <i>ab initio</i> basse résolution des domaines βγ de DivIB de <i>S. pneumoniae</i>	p.112
Figure 41 : Comparaison de l'enveloppe du modèle des domaines βγ de <i>S. pneumoniae</i> (obtenue par modélisation <i>ab initio</i>), avec la structure cristallographique du domaine β d' <i>E. coli</i>	p.112
Figure 42 : Comparaison des courbes de diffusion des différents variants de la partie extracellulaire de DivIB de <i>S. pneumoniae</i>	p.114
Figure 43 : Spectre HSQC [¹⁵ N ; ¹ H] de la partie extracellulaire de DivIB de <i>S. pneumoniae</i> complexée avec le complexe EC/KL.....	p.118
Figure 44 : Superposition des spectres HSQC [¹⁵ N ; ¹ H] du domaine β (¹⁵ N ; ² H) de DivIB de <i>S. pneumoniae</i> , seul et complexé avec EC/KL.....	p.120
Figure 45 : Modèle de la zone d'interaction entre le domaine β et le complexe EC/KL sur la structure du domaine β d' <i>E. coli</i>	p.120

<i>Figure 46</i> : Comparaison de l'interaction avec le complexe EC/KL des protéines $\alpha\beta\gamma$ (ou IB), $\beta\gamma$ et β	p.122
<i>Figure 47</i> : Ajustement des courbes expérimentales d'interaction de la protéine $\beta\gamma$ sur le complexe EC/KL.....	p.124
<i>Figure 48</i> : Séquence primaire des protéines EC, KL, EC* et KL*.....	p.126
<i>Figure 49</i> : Vérification de la formation des complexes EC*/KL et EC/KL*.....	p.126
<i>Figure 50</i> : Comparaison de l'interaction de la protéine $\beta\gamma$ avec le complexe EC/KL, le complexe EC*/KL et le complexe EC/KL*.....	p.126
<i>Figure 51</i> : Mécanismes réactionnels proposés, entre les espèces d'un mélange de complexe EC/KL et de protéine $\beta\gamma$	p.130
<i>Figure 52</i> : Composition d'échantillons équimolaires de mélange de protéine $\beta\gamma$ et de complexe EC/KL, en fonction de la concentration, et de la constante de dissociation K_{D2}	p.132
<i>Figure 53</i> : Courbe de diffusion et fonction de distribution de paires de l'échantillon équimolaire de protéine $\beta\gamma$ et de complexe EC/KL.....	p.132
<i>Figure 54</i> : Modèle <i>ab initio</i> du complexe $\beta\gamma/(EC/KL)_1$	p.134
<i>Figure 55</i> : Comparaison de l'enveloppe du modèle du complexe $\beta\gamma/(EC/KL)_1$ avec le modèle des domaines $\beta\gamma$, obtenus par modélisation <i>ab initio</i>	p.134
<i>Figure 56</i> : Modélisation automatique de corps rigides, de la courbe de diffusion expérimentale du complexe $\beta\gamma/EC/KL$	p.136
<i>Figure 57</i> : Modélisation manuelle de corps rigides, de la courbe de diffusion expérimentale du complexe $\beta\gamma/EC/KL$	p.138
<i>Figure 58</i> : Modèle d'association des protéines EC et KL dans le complexe tétramérique $(EC/KL)_2$	p.163
<i>Figure 59</i> : Comparaisons des séquences primaires et des éléments de structures secondaires des parties extracellulaires de DivIB de <i>G. stearothermophilus</i> et d' <i>E. coli</i> , et comparaison avec la séquence primaire de la partie extracellulaire de DivIB de <i>S. pneumoniae</i>	p.164
<i>Figure 60</i> : Schéma de modèles structuraux des protéines DivIB, DivIC et FtsL, côte à côte dans une membrane.....	p.170
<i>Figure 61</i> : Schéma d'un modèle d'association entre les protéines DivIB, DivIC et FtsL, face à face de part et d'autre de l'invagination, et localisation d'une telle situation sur le schéma d'une bactérie.....	p.172

Liste des tableaux

<i>Tableau 1</i> : Table des longueurs de diffusion cohérente de neutrons de quelques atomes : l'hydrogène, le deutérium, l'isotope 12 du carbone, 14 de l'azote, 16 de l'oxygène.....	p.54
<i>Tableau 2</i> : Valeurs du rayon de Stokes d'échantillons de complexe EC/KL.....	p.76
<i>Tableau 3</i> : Masse apparente du complexe EC/KL déterminée par différentes méthodes.....	p.78
<i>Tableau 4</i> : Valeurs de diffusion à l'origine et de rayon de giration d'échantillons de complexe EC/KL hydrogéné.....	p.86
<i>Tableau 5</i> : Equivalence des intensités à l'origine d'échantillons de complexe EC/KL partiellement deutérés, et complémentaires, dans un tampon 41% D ₂ O/59% H ₂ O.....	p.86
<i>Tableau 6</i> : Valeurs d'intensité de diffusion à l'origine d'échantillons de complexe EC/KL.....	p.92
<i>Tableau 7</i> : Rendement de purification des différents variants de la partie extracellulaire de DivIB de <i>S. pneumoniae</i>	p.102
<i>Tableau 8</i> : Masses apparentes des protéines β , $\beta\gamma$ et $\alpha\beta\gamma$ (ou IB).....	p.102
<i>Tableau 9</i> : Valeurs de diffusion à l'origine et de rayon de giration d'échantillons du domaine β de <i>S. pneumoniae</i>	p.106
<i>Tableau 10</i> : Comparaison des domaines β de <i>S. pneumoniae</i> , <i>E. coli</i> et <i>G. stearothermophilus</i> : taille et rayon de giration.....	p.108
<i>Tableau 11</i> : Valeurs de l'intensité à l'origine et du rayon de giration d'échantillons de la protéine $\beta\gamma$ de <i>S. pneumoniae</i>	p.110
<i>Tableau 12</i> : Valeurs de la diffusion à l'origine et du rayon de giration d'échantillons de la protéine $\alpha\beta\gamma$ (ou IB) de <i>S. pneumoniae</i>	p.114
<i>Tableau 13</i> : Valeurs des constantes de vitesses d'association et de dissociation, des constantes d'association et de dissociation du complexe formé par la protéine $\beta\gamma$ et le complexe EC/KL, et de l'erreur de l'estimation.....	p.124
<i>Tableau 14</i> : Composition des échantillons de mélange de la protéine $\beta\gamma$ et du complexe EC/KL, et leur valeur de diffusion à l'origine.....	p.128
<i>Tableau 15</i> : Calcul du facteur de proportionnalité f_2 , entre l'intensité diffusée à l'origine et le produit de la masse et de la concentration pour les expériences de SAXS.....	p.156

Liste des abréviations utilisées

$1D^1H$: spectre une dimension de l'isotope 1 de l'hydrogène
Å : angström, unité correspondant à 10^{-10} m
Amp : Ampicilline
AUC : Analytical UltraCentrifugation
Da : dalton, unité correspondant à g/mol
DLS : Dynamic Light Scattering
Ec : *Escherichia coli*
EC : protéine constituée du peptide e5 et du domaine extracellulaire de DivIC
EC* : protéine EC tronquée de sa région C-terminale
EC(D) : protéine EC deutérée
EC(H) : protéine EC hydrogénée
EC/KL : complexe formé par les protéines EC et KL
(EC/KL)₁ : complexe formé d'une protéine EC et une protéine KL
(EC/KL)₂ : complexe formé de deux complexes (EC/KL)₁
EDTA: Acide Ethylène-Diamine-Tétraacétique
Gs : *Geobacillus stearothermophilus*
GST : glutathion S-Transferase
Hepes : acide 4-(2-hydroxyéthyl)-1- pipérazine éthane sulfonique
(His6)EC : protéine EC en fusion avec l'étiquette de purification hexahistidine
(His6)EC* : protéine EC* en fusion avec l'étiquette de purification hexahistidine
(His6)KL : protéine KL en fusion avec l'étiquette de purification hexahistidine
(His6)KL* : protéine KL* en fusion avec l'étiquette de purification hexahistidine
HRP : Horse Radish Peroxidase
HSQC : Heteronuclear Single Quantum Coherence
IB : protéine constituée du domaine extracellulaire de DivIB
IPTG : Isopropyl-beta-thio-galactoside
Kan : Kanamycine
KL : protéine constituée du peptide k5 et du domaine extracellulaire de FtsL
KL* : protéine KL tronquée de sa région C-terminale
KL(D) : protéine KL deutérée
KL(H) : protéine KL hydrogénée
LB : Luria broth
M : molaire, unité correspondant à mol/L
NaCl : chlorure de sodium
PBS: Phosphate Buffered Saline
PCR : Polymerase Chain Reaction
PDB : Protein Data Bank
ppm : partie par million
RBM : Rigid body modelling
RMN : Résonance Magnétique Nucléaire
rpm : rotation par minute
RU : Response Unit
SANS : Small Angle Neutron Scattering
SAXS : Small Angle X-ray Scattering
SEC : Size Exclusion Chromatography
SDS-PAGE : Sodium Dodecyl Sulfate PolyAcrylamide Gel Electrophoresis
Sp : *Streptococcus pneumoniae*
SPR : surface plasmon resonance
(Strep)EC : protéine EC en fusion avec l'étiquette de purification Strep
TB : Terrific Broth
TEV : Tobacco Etched Virus
TM : segment transmembranaire
Tris : Trishydroxyméthylaminométhane

Introduction

A. Contexte de l'étude

Sur terre, les animaux et les végétaux n'ont pas le monopole de la vie. Il existe d'autres organismes vivants. Les organismes vivants sont traditionnellement classés en cinq règnes : les animaux, les végétaux, les champignons, les protistes et les procaryotes. Cette classification repose sur la distinction entre les organismes dont l'information génétique est protégée par une membrane, les eucaryotes, à « noyau », des organismes sans noyau, les procaryotes. On sépare les eucaryotes en unicellulaires et pluricellulaires : les protistes et les autres règnes (animaux, végétaux et champignons). Les animaux, les végétaux et les champignons se distinguent ensuite par d'autres règles : les animaux ingèrent leur nourriture, les végétaux la fabriquent, et les champignons se développent dedans. Les procaryotes sont unicellulaires et sont divisés entre deux groupes distincts : les bactéries et les archéobactéries (Campbell, 1995).

I. Des bactéries et des hommes

Les bactéries sont si petites, en moyenne, cent fois plus petites qu'une cellule animale, qu'il est facile d'oublier leur existence. Pourtant ce sont les formes de vie les plus anciennes et, de loin, les plus nombreuses. Une bactérie mesure environ 1 micromètre, et il y aurait environ $5 \cdot 10^{30}$ cellules bactériennes sur terre (Whitman *et al.*, 1998). C'est-à-dire que les bactéries sont $2 \cdot 10^6$ fois plus petites qu'un homme et 10^{21} fois plus nombreuses. On en trouve dans tous les écosystèmes. Elles colonisent même les biotopes les plus extrêmes : les grands fonds marins, les déserts, à plusieurs milliers de mètres sous la surface de la terre... (Rothschild et Mancinelli, 2001). Le biotope que représente les animaux ne fait pas exception : il y a environ dix milliards de cellules bactériennes dans la bouche d'un homme,

mille milliards sur sa peau, un million de milliards dans ses intestins... (Todar, 2007). Cette cohabitation passe la plupart du temps inaperçue. Lorsqu'un des deux partenaires « se nourrit à la table » de l'autre, sans que l'autre s'en rende compte, l'interaction est dite commensale. Cet opportunisme n'est pas parasitaire car l'hôte n'est pas détruit ni même gêné par le profiteur.

Cependant, l'animal peut aussi tirer bénéfice de la présence de bactéries dans son corps. C'est le cas de la flore intestinale par exemple. Dans les intestins de certains animaux, dont l'homme, on trouve plus de 200 espèces de bactéries, et notamment l'espèce *Escherichia coli*. La présence de ces bactéries facilite la digestion car elles dégradent des substances que les animaux ne sont pas capables de dégrader, comme la cellulose, et elles synthétisent des molécules essentielles aux animaux, comme la vitamine K. La présence de bactéries est donc essentielle pour les animaux.

Il arrive aussi que l'équilibre de la cohabitation soit rompu. Des bactéries ne sont pas supportées par leurs hôtes. Elles sont dites pathogènes car elles provoquent des maladies. Un exemple connu de bactérie pathogène est *Yersinia pestis* qui causa la mort de plus de 25 millions d'humains en Europe entre 1347 et 1350. À cette époque, il n'existait pas de médicaments efficaces pour soigner ces maladies infectieuses, et elles étaient alors considérées comme incurables. Une grande découverte a changé cette « fatalité ».

En 1928, Sir Alexander Fleming remarque par hasard qu'un champignon a le pouvoir d'inhiber la prolifération de bactéries (Fleming, 2001). Plus tard, un groupe de chimistes de l'université d'Oxford isole à partir de la même famille de champignons, la molécule qui tue les bactéries et soigne les animaux, la pénicilline (Chain *et al.*, 1940; Ligon, 2004). Les chercheurs découvrent que certains microorganismes ont développé des armes chimiques pour se défendre contre d'autres microorganismes, les molécules antibiotiques. L'homme a donc l'idée d'utiliser ces molécules pour se soigner. Ces molécules « miracles » sont alors fabriquées en grande quantité et utilisées massivement pour guérir les maladies causées par les bactéries et tenter de les éradiquer. Pendant l'âge d'or des antibiotiques (1945-1980), la mortalité baisse, on ne meurt presque plus des maladies bactériennes dans les pays développés.

Mais, les bactéries, comme toutes les organismes vivants, évoluent, et ce, selon le principe de la sélection naturelle (Darwin, 1859). Dès le début de ses recherches sur la pénicilline, Fleming remarque qu'un dosage trop faible, non seulement ne tue pas toutes les bactéries, mais leur permet de développer une résistance à la pénicilline (Fleming, 1945). En

1951, le premier cas de maladie résistante aux traitements antibiotiques est constaté sur un patient atteint de dysenterie, due à *Shigella dysenteriae*. Certaines bactéries peuvent se diviser en une vingtaine de minutes dans des conditions favorables. Toutes les heures, environ deux générations se succèdent. A chaque génération, les individus les plus adaptés aux contraintes du milieu sont naturellement sélectionnés. L'introduction massive des antibiotiques dans la pharmacopée, a agit comme une pression de sélection. Depuis, la résistance des bactéries aux antibiotiques s'est beaucoup développée et est devenue un problème de santé publique, notamment en France, qui détient certains records des taux de résistance aux antibiotiques (BEH, 2006). La science cherche des nouvelles classes d'antibiotiques, auxquelles les bactéries ne sont pas encore résistantes, mais il est important de noter, qu'en théorie, les bactéries s'adapteront à toute nouvelle contrainte, dans un temps plus ou moins long.

II. Un exemple de bactérie commensale, pathogène et résistante : le pneumocoque

Le pneumocoque est une bactérie présente dans les voies respiratoires supérieures d'environ un humain sur deux, de manière asymptomatique (commensale) (Crook *et al.*, 2004). Il est aussi l'agent infectieux de pneumonies et d'otites (pathogène). Plus gravement, il arrive aussi que le pneumocoque traverse la barrière endothéliale des vaisseaux sanguins ou la barrière céphalo-rachidienne, et provoque alors des septicémies et des méningites. Comment le pneumocoque, qui peut être toléré dans la gorge, devient-il pathogène ? Ce n'est pas encore totalement élucidé. En effet, des enfants exposés au pneumocoque peuvent développer une maladie juste après, ou des mois après l'exposition à un pneumocoque. Une personne saine et porteuse d'un pneumocoque, peut transmettre la bactérie à une personne qui, elle, développera une maladie. Le fait que les enfants, les personnes âgées, les personnes immunodéprimées développent plus fréquemment des maladies dues au pneumocoque, indique que celui-ci profite de faiblesses du système immunitaire de l'hôte (Butler, 2004).

Étant un pathogène de l'humain, c'est une bactérie qui a été beaucoup exposée aux antibiotiques, et qui a donc sélectionné des mécanismes de résistance aux antibiotiques. En France, en 2005, 38% des souches isolées de *Streptococcus pneumoniae* étaient de sensibilité réduite à la pénicilline (Gutmann et Varon, 2006).

Le pneumocoque est une bactérie à Gram positif. C'est-à-dire qu'elle fixe la coloration de Gram, technique mise au point pour observer et classer les bactéries en deux groupes, les Gram positif et les Gram négatif. Elle met en évidence les caractéristiques de la paroi bactérienne qui seront plus détaillées au paragraphe suivant. Cette paroi donne sa forme à la bactérie. Le pneumocoque a la forme d'un ellipsoïde allongé, ce qui le classe dans les bactéries ovococoques. Il se divise en plans successifs parallèles, médians à la cellule mère. Lors de sa croissance et lors de la division, les deux cellules filles sont symétriques jusqu'à leur séparation. Le processus de division se déroule en une vingtaine de minutes dans des conditions favorables. C'est un processus continu : les cellules filles commencent à se diviser aussitôt qu'elles sont formées. C'est pourquoi on observe le pneumocoque la plupart du temps sous forme de diplocoques (où les coques des deux cellules filles sont encore soudées) (Figure 1), et parfois en courtes chaînettes (Zapun *et al.*, 2008).

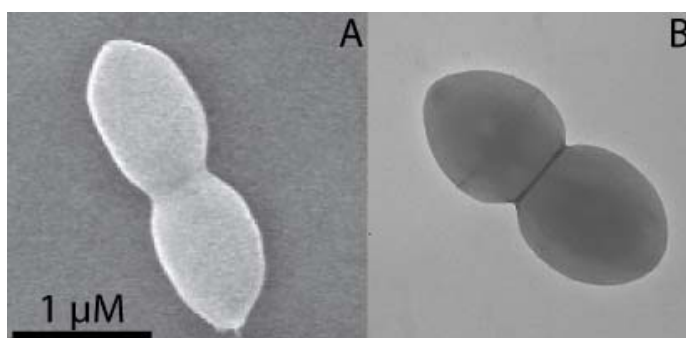


Figure 1 : Images de *S. pneumoniae*, par microscopie électronique. Les images de microscopie électronique à balayage (A) et en transmission (B) ont été réalisées par Anne Marie Di Guilmi (LIM/IBS).

III. La paroi bactérienne

La paroi des bactéries à Gram positif est composée d'une couche épaisse d'un polymère, le peptidoglycane qui entoure la membrane plasmique. Chez une bactérie à Gram négatif, la couche de peptidoglycane est plus mince et est entourée d'une deuxième membrane (Figure 2). La membrane plasmique sépare le cytoplasme qui contient le matériel génétique et le système métabolique, du milieu extérieur. Elle est composée de lipides amphipatiques et contient beaucoup de protéines, dites alors membranaires. La membrane plasmique est une structure flexible et dynamique : les protéines sont mobiles latéralement dans la membrane.

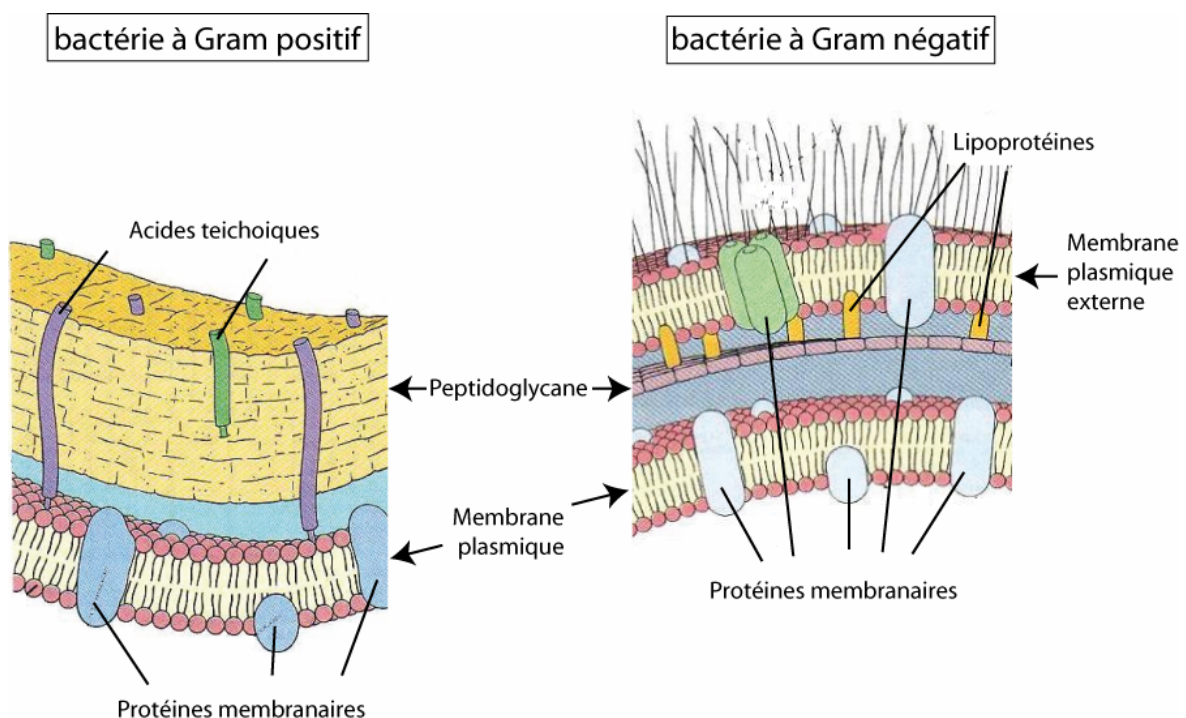


Figure 2 : La paroi bactérienne des bactéries à Gram négatif et à Gram positif. Le cytoplasme d'une bactérie à Gram négatif est isolé par une membrane plasmique, puis par une couche épaisse de peptidoglycane, contenant des acides teichoïques. Les bactéries à Gram positif possèdent une membrane externe maintenue par des lipoprotéines. Les membranes sont composées de phospholipides, en bicouche avec leurs têtes polaires exposées, et de protéines membranaires (Campbell, 1995).

Le peptidoglycane est un polymère de peptides et de sucres, qui protège la bactérie, en agissant comme un exosquelette qui confère sa morphologie à la bactérie. Les chaînes de peptidoglycane sont reliées entre elles par des liaisons peptidiques. Ces pontages font de ce polymère réticulé un réseau élastique qui permet à la bactérie de résister à des pressions

osmotiques importantes. C'est cette protection que cible la pénicilline (Blumberg et Strominger, 1972). Les protéines qui catalysent les réactions de pontage de cette barrière se lient à la pénicilline, c'est pourquoi on les appelle les « penicillin binding proteins » (PBPs) (Tipper et Strominger, 1965). La pénicilline bloque la synthèse du peptidoglycane par l'inhibition des PBPs, et provoque ainsi l'arrêt de croissance et/ou la lyse osmotique des bactéries.

La synthèse du peptidoglycane doit être réalisée sans introduire de discontinuités, pour que l'intégrité de la bactérie soit protégée de la lyse osmotique. Elle se fait par l'insertion d'un précurseur, le lipide II, au peptidoglycane existant (Izaki *et al.*, 1968). Le lipide II est le résultat d'une synthèse enzymatique de plusieurs étapes, impliquant différentes enzymes et substrats (van Heijenoort, 2007). C'est une molécule composée d'un lipide, de sucres et d'une chaîne peptidique. Cette molécule est insérée dans le peptidoglycane par la création de liaisons covalentes, peptidiques et glycanes. La création de ses liaisons est catalysée par des activités enzymatiques, transpeptidase et glycosyltransférase, portées par les PBPs. La synthèse du peptidoglycane est un processus complexe qui n'est que très brièvement abordé ici. Au lecteur intéressé, une revue très complète de T. den Blaauwen est proposée (den Blaauwen *et al.*, 2008).

IV. La division bactérienne

La division d'une cellule bactérienne est un phénomène complexe, qui peut être vu comme la coordination de plusieurs processus. Tout d'abord, la séparation en deux cellules filles nécessite la constriction de la membrane plasmique jusqu'à la fusion qui conduit à la séparation des deux cytoplasmes. Cette constriction doit convertir de l'énergie chimique en énergie mécanique. Elle doit également s'accompagner de synthèse de peptidoglycane. D'autre part, deux cellules filles sont identiquement générées si elles contiennent la même information génétique. Pour cela, l'ADN de la cellule-mère doit être répliqué, et les deux chromosomes transportés, de part et d'autre de la paroi médiane, le septum. Ce transport nécessite aussi une conversion d'énergie et doit être réalisé avant la scission complète. Il est donc évident que ces processus doivent être coordonnés dans le temps, mais aussi dans l'espace, au site de division (Bramhill, 1997).

Des études de génétique ont permis de découvrir un certain nombre de protéines impliquées dans ces processus. Les gènes codant pour ces protéines ont été découverts par mutagenèse des bactéries *E. coli* et *B. subtilis*. Après exposition à un agent chimique (ethyl méthane sulfonate, nitrosoguanidine) provoquant des mutations dans leur génome, certaines bactéries ont présenté un phénotype filamenteux, au dessus d'une certaine température. La filamentation est caractéristique d'un défaut de division : la croissance des bactéries n'est pas altérée, mais seulement leur division. Les gènes mutés dans ces souches filamenteuses sont donc impliqués dans la division. Ces gènes sont appelés *Fts*, pour Filamentous ThermoSensitive (Hirota *et al.*, 1968; Mendelson et Cole, 1972). Le critère de thermosensibilité de ces mutations a été utilisé pour que l'effet de ces mutations ne soit sensible que dans des conditions contrôlées, dans ce cas, la température. Ainsi, les bactéries peuvent être cultivées sans problème à une température dite « permissive », et le phénotype de filamentation induit et étudié à température « non permissive ». On dit que ces souches contiennent des allèles conditionnels de ces gènes.

Ces gènes sont présents chez la plupart des espèces de bactéries. Certains gènes ne possèdent pas la même appellation : les noms « *Fts* » sont attribués aux protéines découvertes chez *E. coli*, bactérie à Gram négatif. Aux protéines découvertes chez les Gram positifs, comme *B. subtilis* et *S. pneumoniae*, on attribue l'appellation « *Div* ». Ainsi, il arrive que des gènes aient deux noms, une pour les Gram positifs, l'autre pour les Gram négatifs, bien qu'ils soient considérés comme orthologues.

Huit de ces protéines sont encodées dans le génome de la plupart des espèces de bactéries (Gram positifs et Gram négatifs) et sont généralement essentielles à leur survie. Ces protéines sont FtsZ, FtsA, FtsK, FtsL, DivIC/FtsB, DivIB/FtsQ, FtsW, FtsI. Pour plus de clarté, dans la suite de ce manuscrit, une seule appellation sera utilisée pour les protéines ayant deux noms, la première citée.

Toutes ces protéines n'ont pas la même fonction dans le processus de division et n'interviennent pas simultanément. Leur recrutement au site de division est séquentiel et conditionnel : le recrutement préalable de certaines de ces protéines en amont permet le recrutement de protéines en aval. Un ordre de recrutement a été proposé pour deux espèces de bactéries, *E. coli* et *B. subtilis* (Errington *et al.*, 2003) (Figure 3). Il est présumé que l'ordre de recrutement chez *S. pneumoniae* soit similaire.

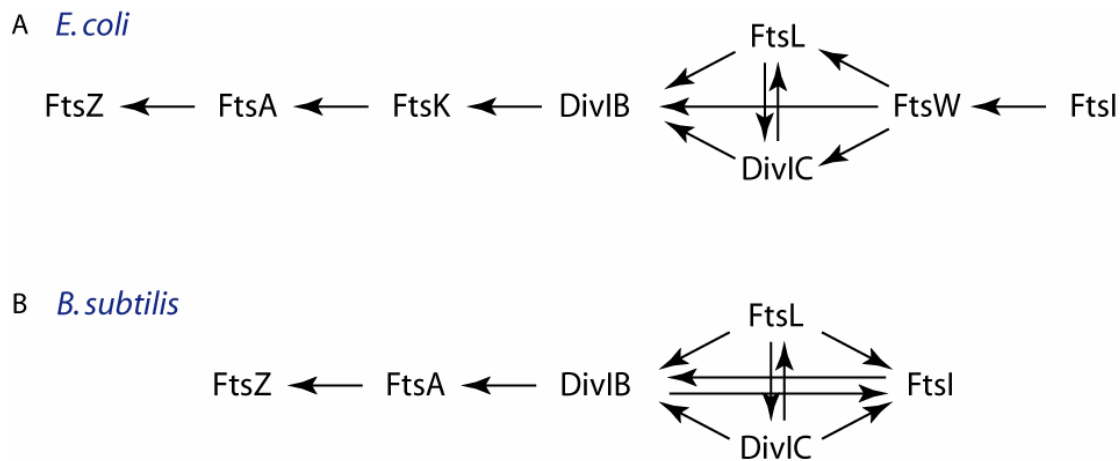


Figure 3 : **Ordre de recrutement de protéines impliquées dans la division, au site de division chez *E. coli* (A), et *B. subtilis* (B)** (d'après Errington *et al.*, 2003).

IV.1. La constriction de la membrane

La protéine impliquée dans le processus de constriction de la membrane est FtsZ. Cette protéine est essentielle chez *E. coli* (Dai et Lutkenhaus, 1991) et *S. pneumoniae* (Margalit *et al.*, 2004), et est présente chez la plupart des bactéries (Margolin, 2000). La microscopie électronique a permis de détecter sa présence au site de division grâce une technique d'immuno-marquage utilisant des billes d'or couplées à des anticorps (Bi et Lutkenhaus, 1991). Cette observation a permis d'émettre l'idée que FtsZ serait impliquée dans l'invagination de la membrane (Lutkenhaus, 1993). D'autre part, cette protéine est capable de polymériser en présence de GTP (Bramhill et Thompson, 1994). Il avait été montré un peu plus tôt qu'elle était une GTPase (de Boer *et al.*, 1992). Par microscopie à fluorescence, des anneaux de FtsZ (exprimée en fusion avec la protéine GFP) ont été observés au site de division de bactéries *E. coli* (Ma *et al.*, 1996). Deux modèles, où la constriction d'anneaux de FtsZ au site de division serait la base du mécanisme d'invagination de la membrane, sont proposés (Bramhill, 1997). Le premier se base sur le glissement de protofilaments de FtsZ généré par un moteur moléculaire non identifié, le second sur la dépolymérisation progressive de protofilaments. La structure de FtsZ, résolue par cristallographie (Lowe et Amos, 1998), montre en effet une grande ressemblance avec la protéine eucaryote tubuline, qui forme des protofilaments et des microtubules. Cette ressemblance structurale était déjà soupçonnée du fait de la similarité de séquence primaire entre FtsZ et la tubuline, et de sa capacité à polymériser avec du GTP (Erickson, 1995).

Les deux modèles proposés exigent que les anneaux de FtsZ soient ancrés à la membrane pour permettre son invagination. C'est l'interaction avec la protéine FtsA qui permettrait cet ancrage. La protéine FtsA est essentielle chez *E. coli* (Lutkenhaus JF, 1979) et *S. pneumoniae* (Lara *et al.*, 2005) et est présente chez la plupart des bactéries (Margolin, 2000). Dans des souches où FtsA n'est pas fonctionnelle, l'anneau de FtsZ se forme, mais l'invagination de la membrane ne se produit pas (Addinall *et al.*, 1996). D'une part, on observe par immunofluorescence la colocalisation de FtsA avec FtsZ, au site de division, mais aussi dans des souches d'*E. coli* où des mutations induisent une localisation différente et non fonctionnelle de FtsZ (Addinall et Lutkenhaus, 1996). Cette colocalisation suggère fortement une interaction. Cette interaction se fait via quelques acides aminés très conservés du domaine C-terminal de FtsZ (Din *et al.*, 1998; Ma et Margolin, 1999; Yan *et al.*, 2000). D'autre part, FtsA possède une hélice amphipatique qui lui permet de s'ancrer dans la membrane (Pichoff et Lutkenhaus, 2005). C'est donc FtsA qui ferait le lien entre l'anneau de FtsZ et la membrane.

Comme FtsA possède un site de fixation de l'ATP (Bork *et al.*, 1992), il était supposé qu'elle puisse fournir l'énergie chimique nécessaire à la constriction mécanique des anneaux de FtsZ. Mais des études ont montré que FtsA n'est capable de se lier à de l'ATP que lorsqu'elle est cytoplasmique ; à la membrane, elle ne peut plus se lier à l'ATP (Sanchez *et al.*, 1994). En ce qui concerne le mécanisme de constriction, un nouveau modèle a été proposé qui se base sur FtsZ, sans apport d'énergie de FtsA (Li *et al.*, 2007). Les progrès de la microscopie électronique ont permis d'observer plus précisément la répartition de FtsZ au site de division que les observations en microscopie à fluorescence ne le permettaient. FtsZ serait réparti en filaments courts et non en filaments continus et longs. Le modèle de constriction de la membrane proposé, repose sur le fait que des filaments de FtsZ se courbent grâce à l'hydrolyse de GTP (Lu *et al.*, 2000). Ainsi, les filaments se fixeraient sur la membrane (via FtsA), se courberaient grâce à l'énergie de molécules de GTP, contraignant la membrane à se courber et donc à diminuer la circonférence de l'anneau de division (Figure 4). Plusieurs itérations successives de ce mécanisme permettraient la constriction progressive de la membrane.

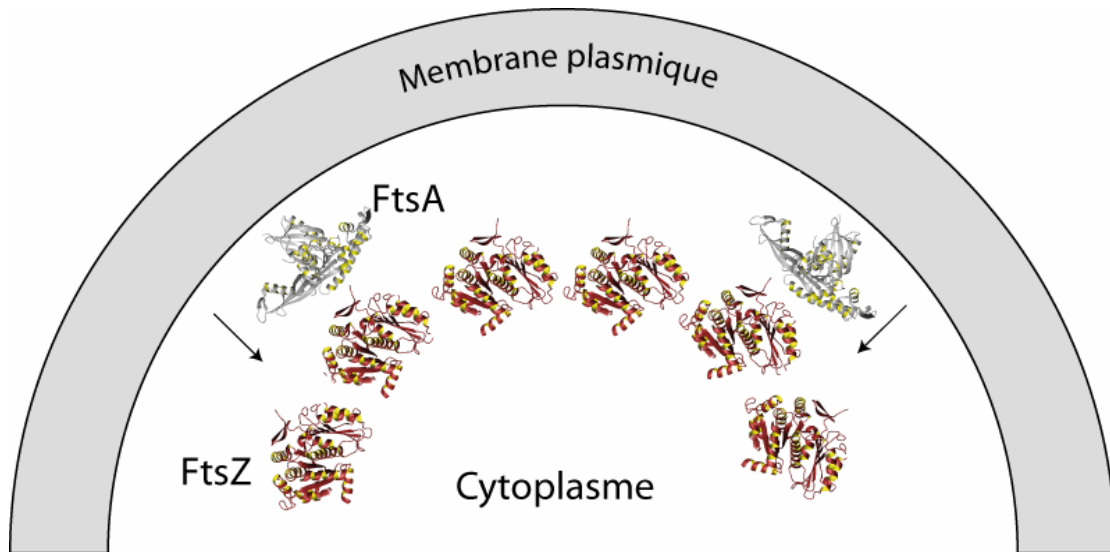


Figure 4 : Assemblage des protéines FtsZ et FtsA à la membrane plasmique. FtsA, en gris, est ancrée à la membrane par une hélice C-terminale qui n'est pas représentée dans la structure (code PDB 1e4g). FtsZ, en brun, (code PDB : 1fsz) interagit avec FtsA par une séquence d'acides aminés en C-terminal, non représentée dans la structure. Le C-ter de FtsZ est orienté vers FtsA. Le C-ter de FtsA est orienté vers la membrane. La courbure de la membrane est contrainte par une courbure du polymère de FtsZ qui entraîne la membrane via FtsA.

IV.2. Le transport des chromosomes

Dans les souches d'*E. coli* où l'allèle de *FtsK* est conditionné par la température, l'inactivation de FtsK provoque des défauts de division, et plus précisément, des défauts de répartition de l'ADN entre les cellules filles (Liu *et al.*, 1998; Yu *et al.*, 1998). Cette protéine est essentielle à la survie d'*E. coli* (Liu *et al.*, 1998) et est présente chez la plupart des bactéries (Margolin, 2000). Dans les souches sauvages d'*E. coli*, FtsK se localise au site de division (Wang et Lutkenhaus, 1998). Une grande similarité de séquence avec une protéine impliquée dans le transport de l'ADN au cours de la sporulation (division particulière, non symétrique) de *B. subtilis*, suggère que FtsK est aussi impliquée dans un mécanisme de translocation d'ADN. En effet, son rôle comme moteur déplaçant de l'ADN par hydrolyse d'ATP a été démontré, par des méthodes biochimiques dans un premier temps (Aussel *et al.*, 2002). Ce moteur a ensuite été analysé, biophysiquement, avec des méthodes de manipulation de molécules uniques (Pease *et al.*, 2005; Saleh *et al.*, 2004). Ces études ont permis d'extraire des informations énergétiques et cinétiques sur ce moteur moléculaire. FtsK peut translater environ 7000 bases d'ADN en une seconde, ce qui en fait le plus rapide moteur moléculaire connu (Strick et Quessada-Vial, 2006).

La résolution de la structure quaternaire de FtsK, par microscopie électronique et ultracentrifugation analytique (Massey *et al.*, 2006) montre que six molécules de FtsK s'assemblent pour former un canal. Ce canal est suffisamment large pour laisser passer un double brin d'ADN (Figure 5). La résolution par cristallographie d'un monomère, puis de l'hexamère de FtsK a permis d'orienter la molécule dans le complexe et de proposer un modèle de fonctionnement pour ce moteur. C'est un moteur qui utiliserait les molécules de FtsK en couple, l'une en face de l'autre. L'hydrolyse d'une molécule d'ATP provoquerait un changement conformationnel, modifiant l'orientation d'un domaine par rapport à un autre domaine, ce qui entraînerait le double brin d'ADN, dans un petit mouvement de translation de 1,6 bases d'ADN. La translation s'accompagnerait d'une petite rotation due à la structure hélicale de l'ADN, qui amènerait le double brin à interagir avec le couple de molécules de FtsK adjacent. Les six molécules de FtsK se relaieraient pour faire avancer le double brin d'ADN, deux par deux. Un couple de FtsK se charge de l'hélice d'ADN pendant que les deux autres couples régénèrent leur ATP. Le mouvement de l'ADN serait donc unidirectionnel, grâce à la structure hélicale de l'ADN.

Ce type de moteur coopératif basé sur six sous-unités est assez conservé en ce qui concerne le transport d'ADN. On le retrouve chez les hélicases, qui séparent les brins d'ADN

(Singleton *et al.*, 2000), et chez les bactériophages, pour compacter l'ARN viral dans les capsules (Mancini *et al.*, 2004). Dans le cas qui nous intéresse, le transport d'ADN se fait pour séparer deux chromosomes de part et d'autre d'une membrane. Pour cela, FtsK est ancré à la membrane via un domaine transmembranaire composé de plusieurs hélices hydrophobes. Dans le modèle proposé par Massey, FtsK est réparti en étoile sur l'anneau de division et forme le canal au centre de l'anneau, pas encore fermé. Le transport de l'ADN se ferait avant la fusion de la membrane. Burton propose que FtsK forme un pore dans la membrane par lequel l'ADN serait exporté, alors que le processus d'invagination de la membrane serait déjà terminé (Burton *et al.*, 2007). La résolution de la structure du domaine cytoplasmique de FtsK ne peut répondre seule à cette question, car une grande région de liaison sépare le domaine d'ancrage à la membrane et celui qui formerait le canal. Ce segment permet plusieurs orientations du canal par rapport à la membrane. L'acteur du transport des chromosomes et son fonctionnement sont maintenant connus, mais le processus global n'est pas encore totalement élucidé.

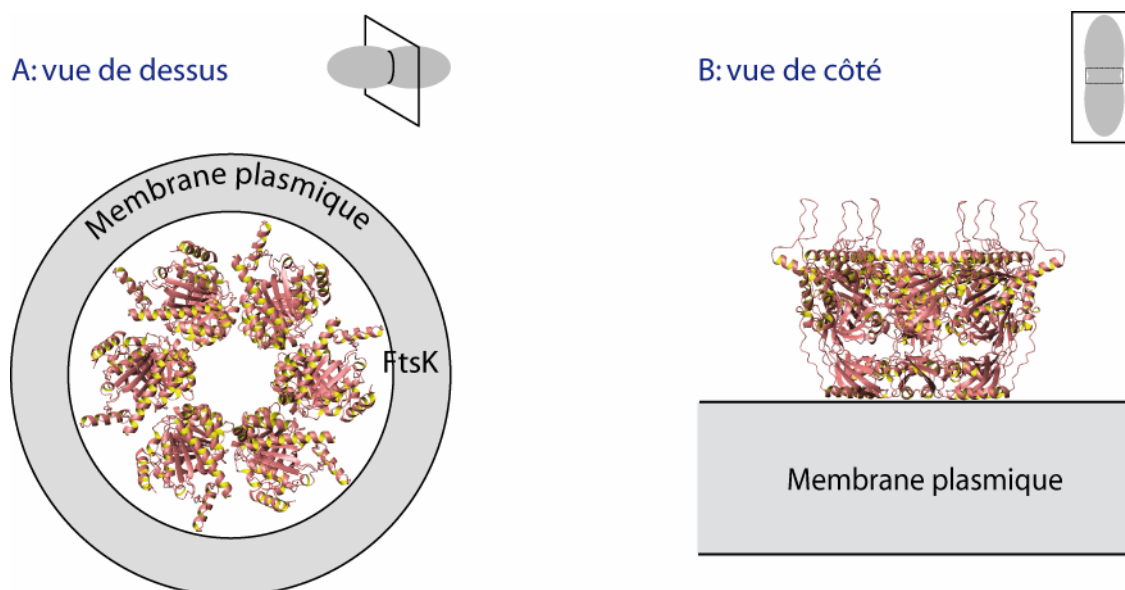


Figure 5 : Le complexe hexamerique de FtsK à la membrane plasmique. Six protéines FtsK, sont ancrées à la membrane par un domaine transmembranaire non représenté dans la structure (code PDB 2iuu) et forment un canal en leur centre pour transporter l'ADN. La vue dite « de dessus » (A) représente une coupe selon le plan de division de la bactérie. La vue dite « de côté » (B) représente une coupe dans le plan perpendiculaire.

IV.3. La synthèse de la paroi septale

De l'autre côté de la membrane a lieu la synthèse de la paroi septale. Plusieurs protéines impliquées dans la synthèse du peptidoglycane ont été découvertes grâce à leur liaison à la pénicilline (Blumberg et Strominger, 1974). L'inhibition de ces protéines par la pénicilline arrête la synthèse de cette paroi. L'inhibition sélective de ces protéines ne provoque pas le même phénotype. La protéine de liaison à la pénicilline 3 (PBP3) d'*E. coli* semble être plus particulièrement impliquée dans la synthèse du peptidoglycane, lors de la division de la bactérie (Spratt, 1975). Cette protéine est codée par le gène *ftsI* (Nishimura *et al.*, 1977), dont des mutations particulières entraînent des défauts de division dans des bactéries qui filamentent (Spratt, 1977). La protéine PBP3/FtsI est donc une protéine essentielle à la division, dont la fonction est de synthétiser le peptidoglycane au septum. Cette protéine est essentielle à la survie d'*E. coli* (Spratt, 1977) et de *S. pneumoniae* (Kell *et al.*, 1993) et est présente chez la plupart des bactéries (Margolin, 2000).

Cette protéine est une enzyme, elle catalyse une réaction chimique de transpeptidation. (Adam *et al.*, 1991; Ishino et Matsubashi, 1981), qui crée une liaison covalente entre deux brins peptidiques du peptidoglycane.

Le précurseur du peptidoglycane, le lipide II est synthétisé dans le cytoplasme. Comme le peptidoglycane et les PBPs se trouvent de l'autre côté de la membrane, il est nécessaire de transporter le lipide II du côté extracellulaire. Une protéine pourrait être impliquée dans ce transport, dans le cadre de la division: FtsW (Ehlert et Holtje, 1996; Ghuysen et Goffin, 1999). Cette protéine membranaire est essentielle à la survie d'*E. coli* (Boyle *et al.*, 1997) et est présente chez la plupart des bactéries (Margolin, 2000). La structure de cette protéine n'est pas encore résolue, mais il est possible que ses dix segments transmembranaires (Gerard *et al.*, 2002) puissent constituer un pore dans la membrane permettant le passage sélectif de molécules de lipide II à travers la membrane. D'autre part, FtsW est lié d'une certaine manière à PBP3/FtsI. D'abord, ces deux protéines sont présentes dans la plupart des espèces de bactéries, et leur présence est très exactement corrélée (Margolin, 2000). Ensuite, les gènes codant pour ces protéines sont transcrits ensemble (Hara *et al.*, 1997). Enfin, FtsW est nécessaire pour la localisation au site de division de FtsI (Mercer et Weiss, 2002). Ces deux protéines semblent donc fonctionner de concert : FtsW transporterait le lipide II du cytoplasme au périplasme, le présenterait à FtsI/PBP3 qui l'incorporerait au peptidoglycane pour former la paroi septale (Figure 6).

Toutefois, ce modèle n'est pas encore validé. En effet, plusieurs points sont encore des hypothèses. Le rôle de FtsW dans le transport du lipide II n'est pas encore avéré et les preuves de la fonction enzymatique de FtsI/PBP3, comme transpeptidase, ne font pas l'unanimité.

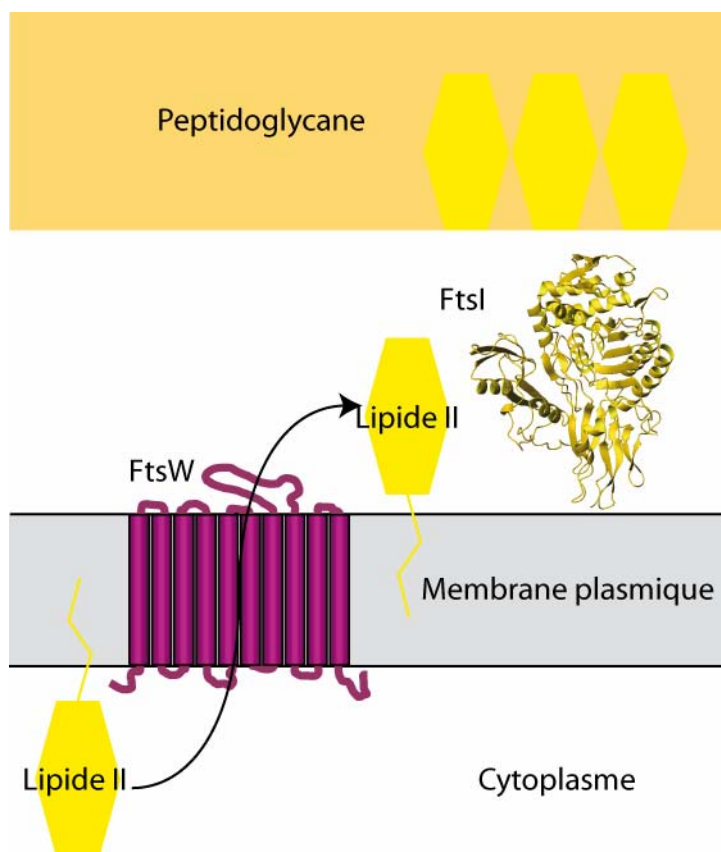


Figure 6 : Assemblage de FtsW et FtsI à la membrane plasmique. La protéine FtsW (en violet) permettrait le transport du lipide II (en jaune vif) du côté externe de la membrane plasmique, pour que FtsI (en jaune, code PDB : 1qme) l'intègre dans le peptidoglycane existant.

V. Trois protéines de la division : DivIB, DivIC et FtsL

Le rôle et le fonctionnement de cinq protéines ont été présentés, parmi les huit citées comme acteurs des étapes majeures de la division. Il reste trois protéines, très conservées et essentielles à la survie de la plupart des bactéries, dont la fonction est peu claire. Ces protéines sont FtsL, DivIC et DivIB.

V.1. FtsL

V.1.1. Quelques données génétiques

FtsL est une protéine essentielle à la survie d'*E. coli* (Guzman *et al.*, 1992), de *B. subtilis* (Daniel *et al.*, 1998) et aussi de *S. pneumoniae* (Le Gouellec *et al.*, 2008). C'est une protéine impliquée dans la division bactérienne puisque des mutations particulières dans le gène codant pour cette protéine provoquent une filamentation des bactéries à une température où l'effet de ces mutations devient sensible. D'autre part elle est localisée au site de division (Ghigo *et al.*, 1999).

Le gène codant pour FtsL se trouve dans une région du chromosome où plusieurs protéines de la division sont codées. C'est pourquoi cette zone a été appelée *dcw* pour « division and cell wall » (Massidda *et al.*, 1998). Chez plusieurs espèces, le gène codant pour FtsL est à proximité du gène codant pour la protéine FtsI. D'ailleurs, chez *B. subtilis*, ces deux gènes sont transcrits ensemble (Daniel *et al.*, 1996), ce qui pourrait suggérer que FtsL soit impliquée dans la synthèse de la paroi septale. La transcription de cet opéron est contrôlée directement par la protéine DnaA, qui est une protéine de la réplication de l'ADN. Le contrôle de transcription d'une protéine de la division par une protéine de la réplication de l'ADN est probablement un moyen de coordonner les processus de la division (Goranov *et al.*, 2005).

V.1.2. Analyse de la séquence primaire

Chez *S. pneumoniae*, le gène *ftsL* code pour une protéine de 105 acides aminés, de 12.2 kDa. La protéine est prédite avec un domaine cytoplasmique court de 25 acides aminés, un segment transmembranaire de la phénylalanine 26 à la leucine 48, et une partie extracellulaire de 57 acides aminés (prédiction réalisée avec le logiciel SOSUI (Hirokawa *et al.*, 1998)) (Figure 7).

<i>S. pneumoniae</i>	1	MAEKMEKT	GQILQMQLKR	FSRVEKAFYF	26	SIAVTTLIVA	ISIIFMQTKL	48
<i>B. subtilis</i>	.MSNLAYQPE		KQQRHAISPE	KKVIVKKRAS	ITLGEKVLIV		LFAAAVLSVS	LLIVSKAYAA	
<i>E. coli</i>	MISRVTEALS		KVKGSMGSHE	RHALPGVIGD	DLRFGKLLPL		CLFICIILTA	VTVVTTAHT	

<i>S. pneumoniae</i>	LQVQNDLTKI	NAQIEEKKTE	LDDAKQEVNE	ILRAEFLKET	A.NSHDLQLN	NENIRIAE	105	.
<i>B. subtilis</i>	YQTNIEVQKL	EEQISSENKQ	IGDLEKSVAD	LSKPCQIMDI	A.KKNGNLNK	DKKVKNIQE	.	.
<i>E. coli</i>	RLLTAQREQL	VLERDALDIE	WRNLILEENA	LGDHSEVERI	ATEKLQMCHV	DPSQENIVVQ	K	

Figure 7 : Alignement des séquences primaires de FtsL de *S. pneumoniae*, *B. subtilis* et *E. coli*. Les identités de séquence sont indiqués par des boîtes rouge. Les acides aminés du segment transmembranaire sont indiqués en gras. Les positions au dessus des acides aminés cités dans le texte sont indiquées, ainsi que la méthionine 1 et le glutamate 105 de FtsL de *S. pneumoniae*. Cet alignement a été réalisé de manière automatique, avec Clustalw.

Dans le domaine extracellulaire, on peut remarquer une répétition d'un motif périodique, composé d'acides aminés hydrophobes (leucine, isoleucine) séparés par des acides aminés chargés ou polaires. Ce type de répétition prédit la formation d'une structure particulière appelé « coiled-coil ». Un « coiled-coil » est l'association longitudinale de deux hélices alpha s'enroulant légèrement l'une autour de l'autre, où les acides aminés hydrophobes des deux hélices sont en regard et en contact, alors que les acides aminés chargés ou polaires sont en contact avec le solvant, et peuvent former des ponts salins entre les chaînes. Les « coiled-coil » sont des structures qui permettent l'interaction entre protéines. Une protéine prédite en « coiled-coil » peut alors interagir, en formant une super hélice de plusieurs hélices alpha, avec elle-même, et constituer des oligomères, comme FtsL semble le faire (Ghigo et Beckwith, 2000), et/ou avec d'autres protéines.

V.1.3. Quelques données biochimiques

FtsL a un temps de demi-vie très court. En effet, chez *B. subtilis*, un arrêt de la synthèse de FtsL provoque très rapidement un arrêt de la division (Daniel *et al.*, 1998). Dix minutes après l'arrêt de la transcription de *ftsL*, il est impossible de détecter la protéine FtsL (Daniel et Errington, 2000). La protéine est donc dégradée très rapidement.

Cette dégradation est sensiblement plus rapide en l'absence de certaines protéines qui sont donc susceptibles d'avoir un effet protecteur de FtsL. En effet, chez *B. subtilis*, dans une souche contenant un allèle conditionnel du gène *divIC*, l'inactivation de DivIC entraîne une disparition de la protéine FtsL (Daniel et Errington, 2000). Chez *E. coli*, l'arrêt de la transcription du gène codant pour DivIC, entraîne aussi une disparition de FtsL (Buddelmeijer *et al.*, 2002). DivIC a donc un effet protecteur sur FtsL. De même, chez *B. subtilis*, dans une

souche contenant un allèle conditionnel de *divIB*, l'inactivation de DivIB entraîne une disparition, plus rapide, de FtsL (Daniel, 2000). Chez *S. pneumoniae*, en absence de DivIB et de synthèse protéique, FtsL est rapidement indétectable (Le Gouellec *et al.*, 2008). DivIB a donc aussi un effet protecteur sur FtsL. Chez *B. subtilis*, l'arrêt de transcription de *ftsI* entraîne aussi une disparition de FtsL. FtsI a aussi un effet protecteur sur FtsL (Daniel *et al.*, 2006). En résumé, trois protéines, DivIC, DivIB et FtsI semble protéger FtsL de la dégradation. Cette protection suggère une interaction entre FtsL et ces protéines. L'interaction entre FtsL et DivIC et DivIB sera abordée dans la section V.5.

Chez *B. subtilis*, la protéase YluC est une des protéase impliquée dans la dégradation de FtsL (Bramkamp *et al.*, 2006). Cette protéase est une protéine membranaire d'une famille de protéases qui clivent les protéines membranaires dans leur segment transmembranaire (Brown *et al.*, 2000). FtsL contient un motif reconnu par YluC dans son domaine cytoplasmique, et est significativement moins dégradée dans des bactéries où le gène *yluC* a été délété. D'ailleurs, la délétion du domaine cytoplasmique de FtsL, donc du motif reconnu par YluC, la protège de la dégradation. Il est intéressant de remarquer que les bactéries où la protéine FtsL a été déléetée de son domaine cytoplasmique sont viables, suggérant que ce domaine de FtsL n'est pas essentiel à son fonctionnement. Il est possible que l'instabilité de FtsL soit importante pour sa fonction. En effet, les bactéries, dont la protéine FtsL a été déléetée de son domaine cytoplasmique, sont plus petites (Bramkamp *et al.*, 2006), ce qui suggère que l'instabilité de FtsL fait partie d'un mécanisme de régulation de la division.

Il a été montré que la transcription du gène codant pour FtsL est contrôlée par une protéine de la réplication de l'ADN, DnaA, et l'arrêt de transcription d'une protéine instable est un moyen d'arrêter le processus dans lequel elle est engagée. FtsL pourrait donc être un point de vérification de bon déroulement de la réplication de l'ADN, nécessaire à la division (Goranov *et al.*, 2005).

V.2. DivIC

DivIC est une protéine essentielle à la survie d'*E. coli* (Ricard et Hirota, 1973), de *B. subtilis* (Levin et Losick, 1994) et de *S. pneumoniae* (Le Gouellec *et al.*, 2008). Comme FtsL, son inactivation entraîne une filamentation des bactéries, et elle se localise au site de division ((Katis *et al.*, 1997), (Noirclerc-Savoye *et al.*, 2005)), ce qui indique qu'elle est impliquée dans la division.

Chez *S. pneumoniae*, le gène *divIC* code pour une protéine de 122 acides aminés, de 14,8 kDa. La protéine est prédite avec un domaine cytoplasmique court de 33 acides aminés, un segment transmembranaire de la méthionine 34 à la tyrosine 56, et une partie extracellulaire de 68 acides aminés (prédiction réalisée avec le logiciel SOSUI (Hirokawa *et al.*, 1998)) (Figure 8).

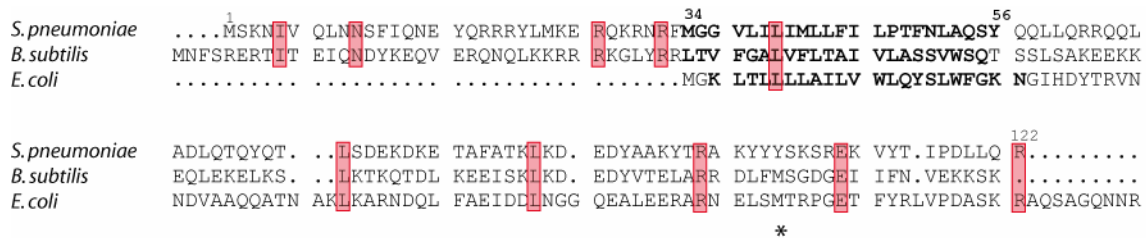


Figure 8 : Alignement des séquences primaires de DivIC de *E. coli*, *B. subtilis* et *S. pneumoniae*. Les identités de séquence sont indiquées par des boîtes rouges. Les acides aminés du segment transmembranaire sont indiqués en gras. Les positions au dessus des acides aminés cités dans le texte sont indiquées, ainsi que la méthionine 1 et l'arginine 122. Une étoile indique la méthionine de la séquence de DivIC de *B. subtilis* citée dans le texte. Cet alignement a été réalisé de manière automatique, avec Clustalw.

Chez *B. subtilis*, la duplication de 5 paires de bases au niveau du codon de la méthionine 110 (Figure 8), entraîne un phénotype filamenteux au dessus d'une certaine température (Levin et Losick, 1994).

Bien que la séquence primaire de DivIC ait une faible similarité entre les espèces bactériennes, un motif de coiled-coil dans le domaine extracellulaire est, lui, très conservé. Ce motif coiled-coil est très similaire à celui de FtsL, et suggère une interaction entre DivIC et FtsL.

V.3. Un complexe entre DivIC et FtsL

Beaucoup d'informations indiquent une interaction entre DivIC et FtsL. Tout d'abord, chez *B. subtilis*, l'arrêt de transcription de l'une entraîne une dégradation rapide de l'autre (Daniel et Errington, 2000; Daniel *et al.*, 1998). Chez *E. coli*, FtsL est indétectable en l'absence de DivIC (Buddelmeijer *et al.*, 2002)). Chez *B. subtilis*, l'une a besoin de l'autre pour se localiser au site de division (Daniel et Errington, 2000; Daniel *et al.*, 1998). Chez *E.*

coli, DivIC a besoin de FtsL pour se localiser (Buddelmeijer *et al.*, 2002). Enfin, un motif « coiled-coil » est présent dans la séquence primaire des deux protéines, et est connu pour médier des interactions entre protéines (Lupas, 1996).

Cette interaction entre FtsL et DivIC a été montrée par la technique de double hybride chez la levure, par migration sur gel non dénaturant (Sievers et Errington, 2000), par co-immunoprécipitation chez *E. coli* (Buddelmeijer et Beckwith, 2004), et par co-purification des protéines de *S. pneumoniae* surexprimées chez *E. coli* (Zapun, non publié). Bien que le « coiled-coil » soit prédit dans le domaine extracellulaire, l'interaction entre les domaines extracellulaires des deux protéines n'a pu être observée (Robson *et al.*, 2002; Zapun, non publié). Les domaines transmembranaires semblent être nécessaires pour stabiliser cette interaction (Buddelmeijer et Beckwith, 2004; Ghigo et Beckwith, 2000; Noirclerc-Savoye *et al.*, 2005). En exprimant le domaine extracellulaire de DivIC, et celui de FtsL, en fusion avec des peptides connus pour former un « coiled-coil », il a été possible de contraindre la formation d'un hétérodimère et de le co-purifier (Noirclerc-Savoye *et al.*, 2005) (cf. Introduction/C.).

Chez *S. coelicolor*, il a été montré que l'inactivation des deux protéines FtsL et DivIC provoquait exactement le même phénotype des bactéries que l'inactivation de l'une ou de l'autre, seule, indiquant que les deux protéines agissent de concert (Bennett *et al.*, 2007). De l'étude de ces mutants, il a été suggéré que le complexe DivIC/FtsL assure la centralisation et symétrisation du septum.

V.4. DivIB

V.4.1. Quelques données génétiques

DivIB est une protéine essentielle à la survie d'*E. coli* (Begg *et al.*, 1980), de *B. subtilis* aux températures élevées (Beall et Lutkenhaus, 1989), mais pas de *S. pneumoniae* (Le Gouellec *et al.*, 2008). C'est une protéine encodée dans le génome de la plupart des bactéries, et notamment de toutes celles qui ont une paroi (Margolin, 2000). L'inactivation de DivIB provoque la filamentation des bactéries et d'autre part, DivIB se localise au site de division (Chen *et al.*, 1999), ce qui indique son implication dans la division bactérienne.

Le gène codant pour cette protéine est, chez la plupart des espèces, localisé juste avant FtsA et FtsZ. Étrangement, chez *S. pneumoniae*, ce gène est éloigné de FtsA et FtsZ, et se trouve proche de gènes *mur*, impliqués dans la synthèse de la paroi (Massidda *et al.*, 1998).

V.4.2. Analyse de la séquence primaire

Chez *S. pneumoniae*, ce gène code pour une protéine de 396 acides aminés, de 45,4 kDa. La protéine est prédite avec un domaine cytoplasmique de 125 acides aminés, un segment transmembranaire de l'isoleucine 126 à la leucine 148, et une partie extracellulaire de 248 acides aminés (prédiction réalisée avec le logiciel SOSUI (Hirokawa *et al.*, 1998)) (Figure 9).

<i>S. pneumoniae</i>	1	MSKDKKNEDK	ETLEELKELS	EWQQRNQEYL	KKKAEEEVAL	AEEKEKERQA	RMGEESEKSE	DKQDQSESTD
<i>B. subtilis</i>	
<i>E. coli</i>	
<i>S. pneumoniae</i>		QEDSESAKEE	SEEKVASSEA	DKEKEEPESK	EKEEQDKKLA	KKATKEKPAK	AKIPGTHIR	AFTILFPSLL
<i>B. subtilis</i>	MNPQDR	EKIVNIEERI	PKIKEQRKQK	AN...RRII	SFIMLFFIMV
<i>E. coli</i>	MS	QAALNTRNSE	EEVSSRRNNG	TRLAGTLELL	TVLTTVLVSG
<i>S. pneumoniae</i>	148	LLIVSAYLLS	PYATMKDIRV	ESTVQTTADD	IRQASGIQDS	DYTINLLLDK	AKYEKQIKSN	Y.WVESAQLV
<i>B. subtilis</i>		LIIV..YLQT	PISKVSTISV	TGNENVSKKE	IIDLSDINSG	D.TEFWSLDK	QKTEKKIQON	K.LVKKAEIS
<i>E. coli</i>		WVVLGWMEDA	QRLPLSKLVL	TSEHYTRND	DIRQSILALG	EPGTFMTQDV	NIIQTQLEQR	LPWIKQVSVR
<i>S. pneumoniae</i>		YQFETKFTIK	VKEVDIVAYY	ISGENHYPII	SSGQLETSSV	SLNSLPETYL	SVLFNDSEKI	KVVFSELAQI
<i>B. subtilis</i>		KSLFNKINIA	IEEYKAIAYL	EKDDVYYEVL	ENGSVLPNEV	TPDDAGP..I	LVNWTNAKKR	SQMAKQLDAL
<i>E. coli</i>		KQWEDELKIH	LVEYVPIARW	NDQHMVDAEG	NTFSVPPDRT	SKQVLPM..L	YGPEGSANEV	LQGYREMGQM
<i>S. pneumoniae</i>		SPELKAAIQK	VELAPSKVTS	DLIRLTMNDS	DEVLVPLSEM	SKKLPPYYSKI	KPOLSEPSV	DMEAGIYSIT
<i>B. subtilis</i>		SNSLKQSISE	IYYTPVKMDE	NRIKLYMNDG	YVVTASIKTF	ADRMKTYP	ISQLSSNKKG	IIHLEVATYF
<i>E. coli</i>		LAKDRFTLKE	AAMTARRSWQLTLNND	IKLNLGRGDT	MKRLARFVEL	YPVLQQQAQT	DGKR..ISV
<i>S. pneumoniae</i>		VADKLIMEAE	EKAKQEAKA	EKKQEEQKK	QEEESNRNQT	NQRSSRR		
<i>B. subtilis</i>		EEFGKNDKAA	KKEDEN....		
<i>E. coli</i>		DLRYDSGAAY	GWAPLPPEES	TQQQNQAQAE	QQ.....		

Figure 9 : Alignement des séquences primaires de DivIB de *S. pneumoniae*, *B. subtilis* et *E. coli*. Les identités de séquence sont indiquées par des boîtes rouges. Les acides aminés du segment transmembranaire sont indiqués en gras. Les positions au dessus des acides aminés cités dans le texte sont indiquées, ainsi que la méthionine 1 et l'arginine 396. Cet alignement a été réalisé de manière automatique, avec Clustalw.

Cette topologie est conservée, quelle que soit l'espèce de bactérie, mais la longueur des domaines cytoplasmique et extracellulaire est très variable selon les espèces. Dans

l'alignement des séquences primaires de DivIB, on remarque la conservation d'une région extracellulaire de 200 acides aminés. La séquence est cependant très variable et, dans un alignement comprenant un plus grand nombre de séquences de DivIB d'espèces bactériennes différentes, un seul acide aminé est entièrement conservé : le glutamate en position 222 chez *S. pneumoniae* (Figure 9) (Le Gouellec *et al.*, 2008). Il est intéressant de remarquer que, dans la souche mutante thermosensible d'*E. coli ftsQI(Ts)*, la mutation de ce seul acide aminé en une lysine, provoque la filamentation des bactéries au dessus d'une température critique (Begg *et al.*, 1980). L'identité des autres acides aminés semblent moins important : en effet, le gène de DivIB de *S. pneumoniae* complémente partiellement la délétion du gène chez *E. coli* (D'Ulisse *et al.*, 2007).

Dans le domaine extracellulaire, la séquence de la méthionine 155 au glutamate 222 peut être identifiée comme un domaine particulier, le motif POTRA (polypeptide-transport-associated). Ce motif est présent dans des protéines impliquées dans le transport ou la protection d'autres protéines (Sanchez-Pulido *et al.*, 2003).

V.4.3. Analyse structurale

V.4.3.1. Chez *G. stearothermophilus*, bactérie à Gram positif

Des expériences de protéolyse limitée ont montré l'organisation en trois domaines du domaine extracellulaire de la protéine DivIB de *Geobacillus stearothermophilus* (bactérie à Gram positif, phylogénétiquement proche de la bactérie *B. subtilis*). En effet, le domaine extracellulaire de la protéine est partiellement digéré par la trypsine. La séquence d'acides aminés proche de la membrane, est protéolysée, délimitant le premier domaine, α . De même, la séquence d'acide aminé terminale est aussi protéolysée, délimitant le dernier domaine, γ . Seule la région centrale, le domaine β , n'est pas dégradée, et donc protégée de la protéase par sa structure. Cette structure de β_{GS} a été résolue par RMN (Robson et King, 2006) (Figure 10). C'est une structure globulaire légèrement aplatie, d'environ 45 Å de diamètres dans le grand axe, et 35 Å de diamètre dans le petit axe. La séquence de β_{GS} est structurée en trois brins bêta, deux hélices alpha, trois brins bêta et deux hélices alpha. Cette structure peut être vue comme présentant deux faces, une composée des six brins bêta, l'autre composée des quatre hélices alpha.

L'étude structurale du domaine β a révélé la présence de deux isomères en solution, une forme *cis* et une forme *trans* de la proline en position 222 chez *G. stearothermophilus*, correspondant à la sérine 327 de *S. pneumoniae* (Figure 9).

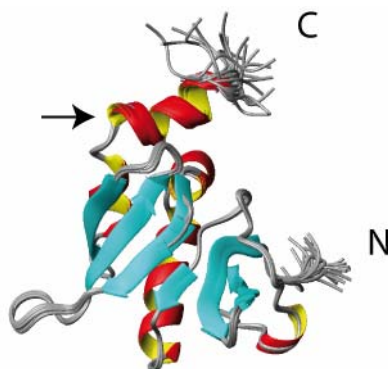


Figure 10 : Structure tridimensionnelle d'une partie de la partie extracellulaire de DivIB de *G. stearothermophilus*. Les extrémités N- et C-terminales de ce domaine, obtenu par protéolyse limitée, sont localisés par les lettres N et C. La position de la proline responsable de deux conformations en solution est indiquée par une flèche. Seule la conformation *trans* est montrée (Robson et King, 2006).

Il a été suggéré qu'un échange conformationnel entre les deux conformères serait utilisé par la bactérie. La commutation entre les deux formes serait un moyen de moduler des interactions avec d'autres protéines. Cependant, cette proline est peu conservée dans les séquences primaires de DivIB des autres espèces de bactérie : elle correspond à une sérine chez *S. pneumoniae*, et à une valine chez *E. coli*, dans l'alignement présenté en figure 9. Cette propriété de DivIB de *G. stearothermophilus* étant peu conservée chez les autres espèces de bactéries, il est permis de douter de la réalité physiologique de ce commutateur moléculaire.

V.4.3.2. Chez *E. coli*, une bactérie à Gram négatif

La résolution de la structure par cristallographie du domaine extracellulaire de DivIB de *E. coli* (bactérie à Gram négatif) a aussi montré une organisation en trois domaines : deux domaines structurés et résolus, α et β , et un dernier domaine γ , non structuré et non résolu (van den Ent *et al.*, 2008) (Figure 11). Les limites de ces domaines sont par contre, légèrement différentes. La limite du domaine α est sensiblement la même (la tyrosine 126 d'*E. coli* correspond au tryptophane 117 de *G. stearothermophilus*). La suite de la séquence est bien structurée jusqu'à la proline 260, délimitant ainsi le domaine β . Un alignement des séquences structurées de DivIB de *G. stearothermophilus* et d'*E. coli*, montre que le domaine bêta de DivIB d'*E. coli* est plus long de 22 acides aminés par rapport à celui de *G. stearothermophilus* (Figure 11).

<i>G. stearothermophilus</i> EWR
<i>E. coli</i>	SKLVLTGERH YTRNDDIRQS ILALGEPGTF MTQDVNIQT QIEQRLPWIK QVSVRKQWPD ELKIHLVEYV
<i>G. stearothermophilus</i>	RIAYVYDRQT FFPLLENGRL LKQEETKTAP SDAPVLVGWK DGDAIAEMTG QLAELPAAVL GAMSEIHYKE
<i>E. coli</i>	PIARWNDQHM VDAEGNTFSV PPERTSKQVL PMLYGPE..G SANEVLQGYR EMGQMLAKDR FTLKEAAMTA
<i>G. stearothermophilus</i>	TREYEDRVIV YMNDGYEVSA TIRQFADKLS HYPATAAALD RNVK.....
<i>E. coli</i>	RRSW...QL TLNNDIKLNL GRGDTMKRLA RFVELYPVLQ QQAQTDGKRI SYVDLRYDSG AAVGWAPLP

Figure 11 : Alignement des séquences primaires des domaines structurés de DivIB de *G. stearothermophilus* et de *E. coli*. Le domaine α est encadré en jaune, le domaine β en orange. Cet alignement a été réalisé de manière automatique, avec Clustalw.

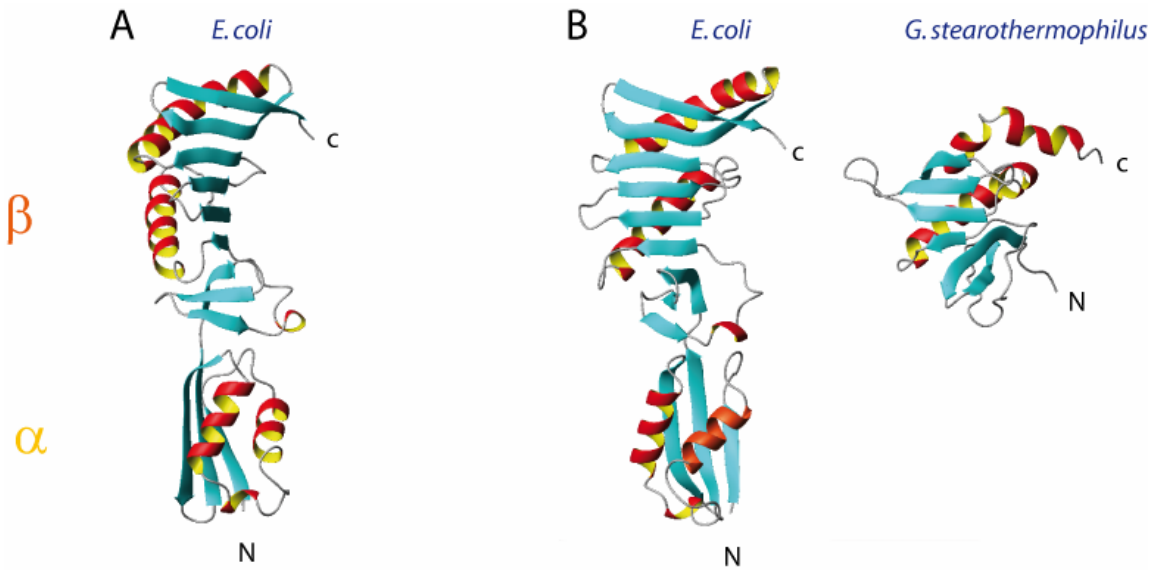


Figure 12 : Structure tridimensionnelle d'une partie du domaine extracellulaire de DivIB de *E. coli* (A) et comparaison avec la structure d'une partie du domaine extracellulaire de *G. stearothermophilus* (B). La structure, (rotation de 90° par rapport à A) est comparée à celle de la protéine isolée par protéolyse limitée de *G. stearothermophilus* (B). Les extrémités N- et C-terminales sont localisées par les lettres N et C. Les domaines distincts de la structure de la protéine d'*E. coli* sont indiqués par les lettres α et β (Robson et King, 2006; van den Ent *et al.*, 2008).

Le domaine extracellulaire de DivIB d'*E. coli*, $\alpha\beta_{EC}$, est allongé, et légèrement courbé (Figure 12). Le domaine α est composé de trois brins bêta et de deux hélices alpha. La topologie de la structure du domaine alpha est un brin bêta, deux hélices alpha et deux brins bêta. Ce domaine présente deux faces, une composée des deux hélices alpha, et l'autre du feuillet bêta. Le domaine β est composé de neuf brins bêta et de deux grandes hélices alpha. Le feuillet bêta compose la face concave de la structure, les deux hélices alpha sont présentées sur la face convexe. La distinction en deux faces (une face feuillet bêta, une face hélice alpha) est plus nette dans la structure β_{EC} .

Les deux structures de β_{EC} et β_{GS} sont comparables. On remarque toutefois l'absence des deux brins bêta C-terminaux (S11 et S12 dans β_{EC}) dans β_{GS} , et une rotation des deux brins bêta N-terminaux (S1 et S2 dans β_{EC}) de la structure β_{GS} par rapport à la structure de β_{EC} (Figure 12).

V.4.4. Quelques données biochimiques et fonctionnelles

La fonction de DivIB n'est pas encore élucidée, cependant plusieurs indices indiquent une possible implication dans la synthèse du peptidoglycane. Tout d'abord, le gène codant pour DivIB n'est présent que dans le génome des bactéries qui possèdent une paroi (Margolin, 2000). D'ailleurs, lorsque des bactéries *E. coli* sont cultivées de manière à les contraindre à perdre leur paroi (dans des milieux isotoniques contenant de la pénicilline), des mutations dans leur génome apparaissent, notamment dans le gène codant pour DivIB. De telles bactéries, appelées protoplastes, ne semblent plus avoir besoin des domaines C-terminaux, β et γ de DivIB puisque ceux-ci ne sont plus codés dans leur génome, le gène étant interrompu par un codon stop (Siddiqui *et al.*, 2006). Ensuite, la sélection de mutant de *B. subtilis* où le gène de DivIB est inactivé, mais dont le phénotype non filamenteux a été restauré, a montré des mutations dans le gène codant pour FtsI. En effet, deux souches de *B. subtilis* ont contourné l'inactivation de DivIB par une mutation génétique. Dans ces souches, un acide aminé de la séquence de FtsI, protéine de la synthèse du peptidoglycane, a été muté (Daniel *et al.*, 2006). Enfin, les souches de *S. pneumoniae* où le gène codant pour DivIB a été délété sont plus sensibles aux antibiotiques de la famille des pénicillines qui inhibent la synthèse du peptidoglycane (Le Gouellec *et al.*, 2008).

Il a aussi été montré que la protéine DivIB a un rôle de protection de la protéine FtsL. Chez *B. subtilis*, FtsL disparaît plus rapidement lorsque la protéine DivIB est inactivée par la température (Daniel et Errington, 2000). Chez *S. pneumoniae*, après l'arrêt de la synthèse protéique, FtsL disparaît beaucoup plus rapidement dans des souches délétées du gène *divIB* (Daniel et Errington, 2000). Il est possible qu'un rôle de DivIB soit de protéger FtsL de la dégradation. En effet, des souches de *S. pneumoniae* délétées du gène codant DivIB sont incapables de pousser en milieu pauvre (milieu CDM), et la surexpression de FtsL dans des ces souches restaure leur croissance (Le Gouellec *et al.*, 2008). D'autre part, l'essentialité de DivIB chez *B. subtilis* à haute température, peut être contournée par la surexpression de FtsL (Daniel et Errington, 2000). Un rôle de DivIB pourrait être de réguler la synthèse du peptidoglycane, via la quantité de FtsL présente au site de division.

Cette protection de FtsL par DivIB indique que ces deux protéines interagissent.

V.5. Un complexe entre FtsL, DivIC et DivIB

Plusieurs indices suggèrent donc que DivIB, FtsL et DivIC interagissent.

Tout d'abord, DivIB protège FtsL de la dégradation chez *B. subtilis* (Daniel et Errington, 2000) et *S. pneumoniae* (Le Gouellec *et al.*, 2008). Chez *B. subtilis*, l'action de DivIB sur DivIC est inverse : DivIC est dégradée plus rapidement quand DivIB est surexprimée (Daniel *et al.*, 2006). De plus, pour se localiser au site de division, FtsL a besoin de DivIB chez *B. subtilis* (Daniel et Errington, 2000) et chez *E. coli* (Ghigo *et al.*, 1999). DivIC a besoin de DivIB chez *E. coli* (Buddelmeijer *et al.*, 2002) et chez *B. subtilis* à haute température (Katis *et al.*, 2000). Chez *B. subtilis*, DivIB a aussi besoin de FtsL et de DivIC pour se localiser (Daniel *et al.*, 1998; Katis *et al.*, 2000).

D'autre part, chez *E. coli*, une co-immunoprécipitation de FtsL et DivIC n'est possible qu'en présence de DivIB (Buddelmeijer et Beckwith, 2004). D'ailleurs, dans le complexe formé et immuno-précipité, DivIB a été détecté par western blot. Enfin, la co-purification du domaine extracellulaire de DivIB de *S. pneumoniae* a été réalisée avec un dimère artificiel et contraint de DivIC et FtsL (Noirclerc-Savoye *et al.*, 2005).

VI. Synthèse

Huit gènes se sont distingués du génome de la plupart des espèces de bactérie par leur implication dans le processus vital de la division. Ces huit gènes sont *FtsZ*, *FtsA*, *FtsK*, *DivIB*, *DivIC*, *FtsL*, *FtsW* et *FtsI*. Les protéines FtsZ et FtsA sont les acteurs du processus de constriction de la membrane plasmique qui va délimiter les deux cellules filles. La protéine FtsK transloque les deux chromosomes répliqués de part et d'autre de cette membrane. Les protéines FtsW et FtsI sont vraisemblablement impliquées dans la synthèse du nouveau peptidoglycane nécessaire au site de division (Figure 13). Par contre, le rôle de DivIB, FtsL et DivIC n'est pas encore totalement éclairci.

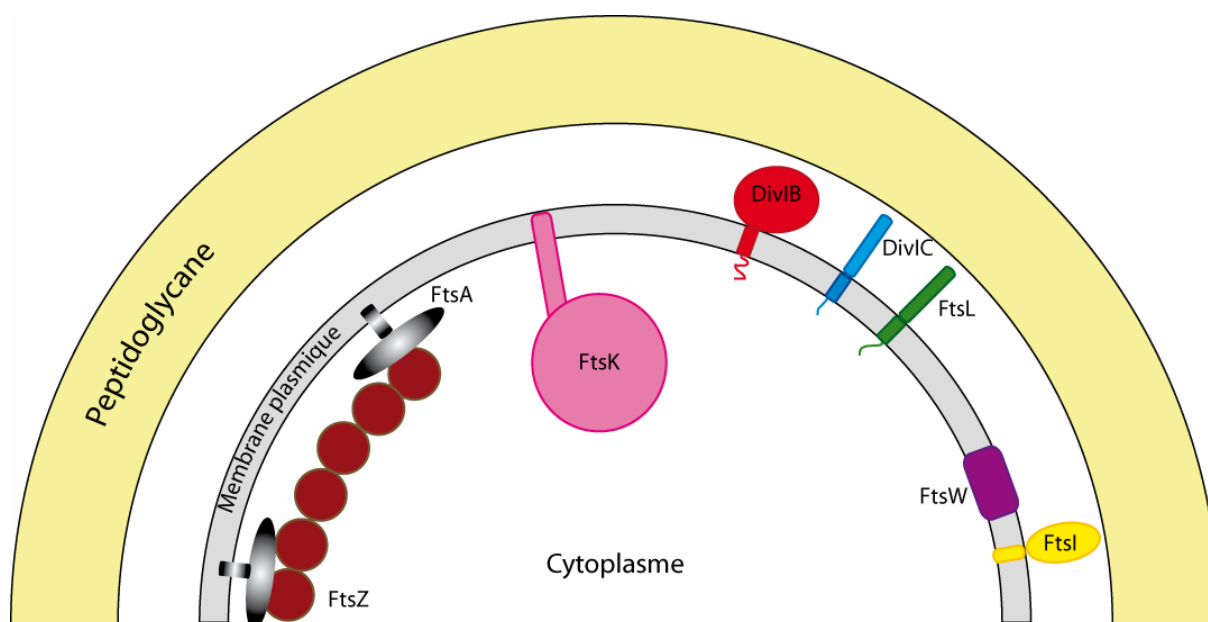


Figure 13: Assemblage des protéines de la division bactérienne. Cette représentation est basée sur les prédictions topologiques des protéines et quelques informations d'association. FtsZ (en brun), forme des filaments qui contraignent la membrane à se courber via FtsA (en dégradé de gris), ancrée à la membrane. FtsK, en rose, possède un grand domaine cytoplasmique qui forme un hexa-complexe. DivIB (en rouge), DivIC (en bleu) et FtsL (en vert) sont majoritairement extracellulaires. FtsW (en violet), quasiment intégralement membranaire, et FtsI (en jaune), principalement extracellulaire, sont impliquées dans la synthèse du peptidoglycane au site de division.

Beaucoup de questions restent ouvertes sur les protéines DivIB, DivIC et FtsL, notamment structurales. En effet, quasiment aucune information structurale n'a encore été apportée sur les protéines DivIC et FtsL, et deux structures, différentes, de domaines de la protéine DivIB existent.

B. Les outils de la physique

La biologie s'intéresse aux phénomènes du vivant, des écosystèmes aux molécules. Avec les molécules, s'ouvre le domaine de la chimie et, avec les atomes, celui de la physique des particules. Bien que ce monde microscopique ne soit pas visible à l'oeil nu, il n'en reste pas moins très complexe. Comment étudier et comprendre ce monde invisible ? La première étape est sûrement de pouvoir le « voir ».

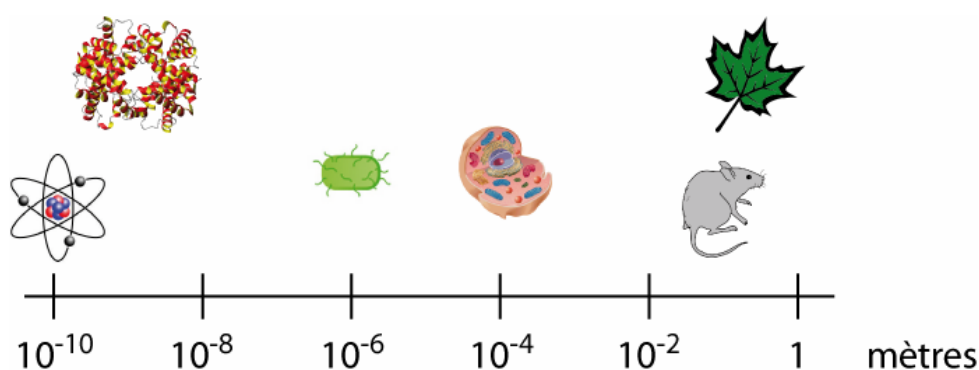


Figure 14 : Quelques ordres de grandeur de la biologie. Sur une échelle logarithmique, sont représentés une feuille et une souris qui mesurent quelques centimètres, une cellule animale qui mesure 0,1 mm environ, une bactérie *E. coli* qui mesure environ un micromètre, la structure de l'hémoglobine (code PDB : 1A00) qui mesure environ 50 Å, et un atome qui mesure environ un angström.

Les quelques méthodes présentées dans la suite, sont celles qui ont été utilisées ou abordées lors de mon travail de recherche, pour tenter d'obtenir des informations structurales sur un système de protéines en interaction (excepté les microscopies, qui ne sont évoquées que de manière introductive).

I. Les microscopies

Grâce aux microscopes optiques, on peut voir certains détails des cellules eucaryotes et assister, par exemple, à la fascinante mitose. On peut également voir les bactéries et observer leur division, mais il est difficile de distinguer leurs composants. En effet, la résolution d'un microscope est limitée par la diffraction de la lumière. Les objets de taille proche des

longueurs d'onde utilisées diffractent les rayons lumineux. Du fait de cette diffraction, la distance de résolution R , c'est-à-dire la distance minimale à laquelle deux objets sont vus séparément, vaut:

Eq. 1 :

$$R \approx 0,61 \frac{\lambda}{n \cdot \sin u}$$

où λ est la longueur d'onde, n est l'indice de réfraction du milieu, u est la moitié de l'angle sous lequel on voit l'objectif du microscope depuis l'objet (considéré comme ponctuel) (Serdyuck *et al.*, 2007). La quantité $n \cdot \sin u$ s'appelle l'ouverture numérique, et est un paramètre important de tout microscope. Les ondes de la lumière visible ont une longueur allant de 0,4 μm à 0,8 μm , limitant la résolution des microscopes optiques à environ 1 μm .

La limite de résolution des microscopes électroniques est plus petite puisqu'ils reposent sur des rayons à une autre longueur d'onde. Dans les microscopes électroniques, ce ne sont pas des photons mais des électrons qui transportent l'information. Ces électrons qui sont des particules avec une quantité de mouvement \vec{p} , peuvent aussi être considérés comme des ondes, de longueur d'onde fixée par l'équation de De Broglie $\lambda = \frac{h}{p}$, où h est la constante

de Planck. Ces ondes électrons interagissent avec le potentiel électrostatique des atomes de l'échantillon et sont ensuite refocalisées par des lentilles magnétiques. Le principe est assez comparable à celui des microscopes optiques : des lentilles refocalisent les rayons diffusés par l'objet éclairé, et l'image est ensuite reformée sur un plan où l'on place un détecteur. En théorie la résolution est fixée par la longueur de l'onde avec laquelle on éclaire l'objet, mais pour ce type de microscope, elle est surtout limitée par la qualité des lentilles et par la dégradation de l'échantillon. En effet, les lentilles magnétiques, qui sont des champs magnétiques appliqués aux électrons pour les refocaliser, ne sont pas aussi efficaces que les lentilles optiques : la résolution est limitée par les aberrations géométriques qu'elles introduisent dans l'image. D'autre part, les électrons sont des particules ionisantes qui dégradent les échantillons. Cette dégradation introduit un bruit dans le signal détecté et limite ainsi la résolution. Ces contraintes de lentilles et de dégradation gardent invisibles, pour l'instant, les molécules biologiques d'une taille inférieure à 500 kDa, et lorsque les objets sont visibles, il est possible d'atteindre une résolution de l'image de l'ordre de 5 angströms.

II. La diffusion/diffraction

Pour visualiser un objet, il faut l'éclairer avec une onde dont la longueur d'onde est du même ordre de grandeur que la taille de l'objet. Donc pour regarder les objets trop petits pour être distingués avec un microscope électronique, on utilise des photons X ou des neutrons, dont la longueur d'onde est de l'ordre de l'angström. Le problème est que pour ces particules, il n'existe pas encore de lentilles capables de refocaliser les rayons, comme les lentilles optiques pour les photons visibles ou les lentilles magnétiques pour les électrons, pour former une image.

Il est cependant possible de le faire mathématiquement. Un assemblage d'atomes, éclairé par un faisceau incident, diffuse une partie de l'onde incidente, dans toutes les directions. L'onde diffusée est la somme des ondes diffusées par chaque atome :

$$\text{Eq. 2 :} \quad \vec{F}(\vec{Q}) = \sum_j f_j \exp(i \cdot \vec{Q} \cdot \vec{r}_j)$$

$$\text{Eq. 3:} \quad Q = \frac{4\pi}{\lambda} \sin\left(\frac{\theta}{2}\right)$$

où f_j est l'amplitude de diffusion de l'atome j , \vec{r}_j son vecteur position, et \vec{Q} le vecteur différence entre l'onde incidente et l'onde diffusée, λ est la longueur d'onde et θ , l'angle entre l'onde diffusée et l'onde incidente (Serdyuck *et al.*, 2007).

La carte de densité $\rho(\vec{r})$ de l'assemblage d'atomes, où \vec{r} est le vecteur coordonnée de chaque atome, est relié à la somme des ondes diffusées $F(\vec{Q})$ par transformée de Fourier. En connaissant $F(\vec{Q})$, on peut calculer $\rho(\vec{r})$ et construire une représentation tridimensionnelle de l'objet. Cette somme, $F(\vec{Q})$ est un nombre complexe : elle contient les informations d'intensité mais aussi de déphasages des ondes diffusées les unes par rapport aux autres. Le problème est que l'on est capable de ne mesurer, que le module de cette somme, l'intensité, mais pas sa phase. Sans les phases des ondes diffusées, on ne connaît que partiellement ces ondes diffusées et le calcul de $\rho(\vec{r})$ est impossible.

Deux approches permettent de contourner ce problème de mesure. La cristallographie et la diffusion aux petits angles.

II.1. La cristallographie

Dans les années 50, des biophysiciens ont eu l'idée d'appliquer aux molécules biologiques, les méthodes de cristallographie. En 1958, J.C. Kendrew élucide la structure moléculaire de la myoglobine grâce à des clichés de diffraction de cristaux de cette molécule (Kendrew *et al.*, 1958). En 1960, c'est la structure de l'hémoglobine qui est résolue par M.F. Perutz (Perutz, 1960). A ce jour, on connaît la structure de plus de 39 000 molécules biologiques (protéines, acides nucléiques...) grâce à la diffraction de rayons X.

En traversant un réseau ordonné de molécules biologiques, un faisceau de rayons X est diffracté, c'est-à-dire éclaté, dans des directions particulières liées aux dimensions du réseau cristallin. On mesure l'intensité diffractée en fonction de la direction,

Eq. 4 :

$$I(\vec{Q}) = |\vec{F}(\vec{Q})|^2$$

où \vec{Q} est la différence entre le vecteur du faisceau diffracté et celui du faisceau incident. L'analyse de $I(\vec{Q})$, permet de résoudre le groupe d'espace du cristal, on trouve ainsi les paramètres du réseau cristallin. Pour connaître le déphasage des ondes diffusées lié aux positions des atomes dans la molécule, on peut alors utiliser différentes méthodes (remplacement moléculaire, remplacement multiple isomorphe, dispersion anormale).

Lorsqu'on obtient ces phases des ondes les uns par rapport aux autres, si les jeux de données de diffraction sont de bonne qualité, il est possible de calculer la position des atomes d'une molécule dans l'espace. On dit que la structure de la molécule a été résolue à haute résolution. Il est alors possible, à partir des coordonnées des atomes de construire une représentation tridimensionnelle des molécules.

II.2 La diffusion aux petits angles

Dans l'approche de la diffusion aux petits angles, on se concentre uniquement sur l'intensité pour en extraire le maximum d'informations possibles.

En cristallographie, on considère une situation où la molécule est ordonnée dans un réseau cristallin. Pour la diffusion aux petits angles, on considère une situation inverse : les molécules sont complètement désordonnées et sont dans toutes les orientations possibles. Si l'on considère qu'elles n'interagissent pas du tout entre elles, alors l'intensité diffusée devient la somme des intensités diffusées par chaque molécule. Comme elles prennent toutes les orientations possibles, l'intensité diffusée est alors la somme de l'intensité, moyennée sur toutes les orientations.

Eq. 5 :
$$I(Q) = N \left\langle \left| \vec{F}(\vec{Q}) \right|^2 \right\rangle$$

$I(Q)$ devient une fonction de Q qui n'est plus un vecteur : on ne peut extraire aucune information sur l'orientation car elles sont moyennées. On ne peut extraire que des informations de distance. Dans les expériences de diffusion aux petits angles, l'intensité isotrope ne dépend que de la distance à l'origine (point d'impact du faisceau direct, non diffusé), et plus de son orientation dans l'espace.

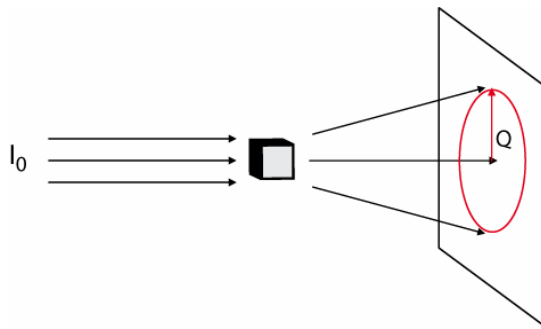


Figure 15 : Représentation schématique de la diffusion d'un faisceau incident I_0 par un objet. L'intensité diffusée $I(Q)$ est représentée sur un détecteur par un cercle rouge.

II.2.1. La formule de Debye

Cette moyenne $I(Q)$ de l'intensité diffusée par un ensemble de molécule dans une solution a été calculée par P. Debye. Elle s'exprime en fonction des distances inter atomiques de la molécule. Cette formule est la base de la technique de la diffusion aux petits angles.

$$\text{Eq. 6 :} \quad I(Q) = N \sum_j \sum_k m_j m_k \frac{\sin(Qr_{jk})}{Qr_{jk}}$$

$$\text{Eq. 7 :} \quad m_j = f_j - \rho^0 v_j$$

où N est le nombre de molécules, Q est le module de la différence entre l'onde diffusée et l'onde incidente, m_j , l'amplitude de diffusion j de la position j par rapport au solvant de densité ρ^0 et de volume v_j , et r_{jk} la distance entre la position j et la position k . Dans cette formule, l'intensité est reliée aux distances entre les atomes de la molécule, et donc aux dimensions de cette molécule. Pour la gamme de $I(Q)$ où le produit $r_{jk}Q$ est petit, donc aux petits angles, la formule peut être développée en série limitée aux deux premiers termes :

$$\text{Eq. 8 :} \quad I(Q) = N \left(\sum_j m_j \right)^2 \left(1 - \frac{R_G^2 Q^2}{3!} \right)$$

$$\text{Eq. 9 :} \quad R_G = \frac{\sum m_j m_k r_{jk}^2}{2 \left(\sum m_j \right)^2}$$

$$\text{Eq. 10 :} \quad I(0) = N \left(\sum_{jk} m_j m_k \right) = \left(\sum m_j \right)^2.$$

$I(0)$ est l'intensité émise pour un angle de diffusion nulle. Elle ne peut être mesurée car elle est masquée par l'intensité du faisceau direct. Elle est donc extrapolée de la courbe $I(Q)$. À l'angle nul, l'intensité diffusée est la somme simple des amplitudes de diffusion, sans facteur dû aux distances entre les centres diffusants. Ainsi, on peut la formuler en fonction de la masse de l'objet qui diffuse, pour une solution monodisperse:

$$\text{Eq. 11 :} \quad I(0) = N \left(\sum m_j \right)^2 = c N_A M \left(\frac{\sum m_j}{M} \right)^2$$

où c est la concentration de l'échantillon en mg/mL, N_A le nombre d'Avogadro, M la masse molaire de la molécule, m_j l'amplitude de diffusion par rapport au solvant de l'atome j .

Cette formulation permet de calculer la masse en solution de l'objet qui diffuse. Elle peut être différente de la masse théorique de la molécule, dans le cas d'un dimère par exemple, où la masse mesurée sera deux fois celle attendue. Cette information est très importante pour connaître le degré d'oligomérisation de l'objet étudié en solution. La technique de la diffusion aux petits angles donne accès à cette information sans modèle initial.

Avec la diffusion aux petits angles de neutrons, il est possible de calibrer de manière absolue cette intensité.

$$\text{Eq. 12 :} \quad I(0)_{\text{exp}} = f_1 \cdot M \cdot c \cdot I(0)_{\text{lampon}}$$

Le facteur de proportionnalité, f_1 , entre l'intensité à l'origine, le produit de la masse molaire et la concentration peut être calculé avec les paramètres de l'expérience (Jacrot et Zaccai, 1981).

Avec la diffusion de rayons X, on peut calibrer cette intensité relativement à l'intensité à l'origine d'un échantillon de masse molaire et de concentration connues (Mylonas et Svergun, 2007).

$$\text{Eq. 13 :} \quad I(0)_{\text{exp}} = f_2 \cdot M \cdot c$$

Le facteur de proportionnalité f_2 est alors déterminé expérimentalement.

Le rayon de giration R_G , de la distribution des amplitudes de diffusion par rapport au solvant, m_j , est une caractéristique de la molécule. Il est lié à sa forme. Pour une masse donnée, le rayon de giration d'un objet est différent selon sa forme. R_G est minimal dans le cas d'une sphère et plus grand dans le cas d'une forme éloignée d'une sphère.

Une fonction gaussienne est proposée par A. Guinier comme meilleure approximation (Guinier et Fournet, 1955), valable dans la gamme des valeurs de Q , où le produit $R_G^2 Q^2$ est inférieur à 1:

$$\text{Eq. 14 :} \quad I(Q) = I(0) \exp \left[-\frac{R_G^2 Q^2}{3} \right]$$

Grâce à cette formulation, l'estimation des paramètres $I(0)$ et R_G est mathématiquement plus simple :

Eq. 15 :

$$\ln I(Q) = \ln I(0) - \frac{R_G^2 Q^2}{3}$$

En représentant $\ln I(Q)$ en fonction de Q^2 , dans la zone où la courbe peut être approximée par une droite (la zone de Guinier), R_G^2 est la pente de cette droite, multiplié par un facteur 3, et $I(0)$, l'ordonnée à l'origine. De plus, cette représentation permet une analyse visuelle rapide des données.

Ainsi, l'étude de la courbe de diffusion $I(Q)$ dans la zone des petits angles permet, dans un premier temps, de connaître la masse M , donc le degré d'oligomérisation de l'objet irradié, et d'en mesurer une dimension caractéristique propre, le rayon de giration R_G .

II.2.2. La fonction de distribution de distance

La courbe de diffusion $I(Q)$ contient de l'information relative à la forme de l'objet irradié, sur toute la gamme de Q , et pas seulement dans la zone de Guinier. La zone des petits angles contient des informations sur les grandes distances de l'objet. Aux plus grands angles, ce sont les distances plus petites qui contribuent à la forme de la courbe. La transformée de Fourier de cette courbe d'intensité moyennée peut être calculée. Sans révéler les coordonnées des atomes, sa forme reflète la distribution des distances entre ces atomes. La fonction calculée ainsi, $p(r)$, s'appelle la fonction de distribution de distances. Elle est analogue à la fonction de Patterson utilisée en cristallographie. Cette fonction est dans l'espace réel, c'est-à-dire qu'elle s'exprime en fonction de r , distance réelle. Cette fonction représente le nombre de fois qu'une distance r , entre deux centres diffuseurs, intervient dans la diffusion. Elle est proportionnelle au nombre de fois que cette distance est présente dans l'objet, et ainsi, la distribution des longueurs dans l'objet. Une sphère ne contient pas la même distribution de longueurs qu'un cylindre.

GNOM est un logiciel qui a été développé pour estimer cette fonction de distribution de distances (Semenyuk et Svergun, 1991). À partir d'une courbe de diffusion, la fonction est calculée par ajustement manuel de la distance maximal de la molécule.

La distribution de distances est caractéristique d'une forme. Grâce à cette fonction de distribution de distances, il est possible d'estimer la forme globale de l'objet diffuseur. DAMMIN est un logiciel qui a été développé pour réaliser des modélisations de l'objet

diffuseur à partir de cette fonction de distribution de distances (Svergun, 1999). Grâce à ces programmes, la diffusion aux petits angles permet, pour des données de qualité suffisante, de proposer un modèle tridimensionnel des molécules étudiées. Ces modèles sont dits à basse résolution car, les distances inférieures à la dizaine d'angströms (dans les meilleurs modèles) ne sont pas résolues. La limite de résolution de cette technique est liée à l'extrême angle Q pour lequel une intensité est mesurée. En effet,

Eq. 16 :

$$Q_{\max} = \frac{2\pi}{d_{\min}}$$

Pour certaines molécules, cette basse résolution est la seule accessible. Les modèles construits apportent une information qui, bien qu'incomplète, permet une visualisation de la forme des molécules.

II.2.3. La complémentarité rayons X/neutrons

Deux types de particules peuvent être utilisés pour la diffusion aux petits angles: les photons X et les neutrons. L'utilisation de ces deux particules n'est pas redondante mais complémentaire. L'onde diffusée par un assemblage d'atome est, quelle que soit la nature du faisceau incident :

Eq. 17 :

$$\vec{F}(\vec{Q}) = \sum_j f_j \exp(i\vec{Q}\vec{r}_j)$$

La différence entre les photons X et les neutrons est dans l'amplitude de diffusion f_j . Les photons interagissent avec et sont diffusés par le nuage électronique des atomes. L'efficacité d'interaction est simplement proportionnelle à la taille de ce nuage électronique. L'amplitude de diffusion d'un électron a été calculée par J.J. Thomson et s'appelle le facteur de Thomson, f_{e1} . Le facteur de diffusion d'un atome à n électrons est donc $n f_{e1}$.

Les neutrons interagissent avec les noyaux. L'interaction n'est pas plus importante avec la taille des noyaux, et peut être très différente pour deux isotopes. Les valeurs de l'amplitude de diffusion pour les différents noyaux sont tabulées.

atome	^1H	$^2\text{H(D)}$	^{12}C	^{14}N	^{16}O
$b_{\text{coh}} (10^{-14} \text{ m})$	-37,4	66,7	66,5	93,6	58,0

Tableau 1 : Table des longueurs de diffusion cohérente de quelques atomes : l'hydrogène, le deutérium, l'isotope 12 du carbone, 14 de l'azote, 16 de l'oxygène (Sears, 1992).

Il est intéressant de remarquer que cette amplitude est négative pour l'hydrogène, et très différente d'un de ses isotopes, le deutérium. Cette différence permet une utilisation astucieuse des neutrons.

En effet, la plupart du temps, une molécule est étudiée en solution aqueuse. La diffusion de cette molécule s'exprime alors par rapport à la diffusion de la solution. L'onde diffusée par la molécule devient

Eq. 18 :

$$\vec{F}(\vec{Q}) = \sum_j (f_j - \rho^\circ v_j) \exp(i\vec{Q}\vec{r}_j)$$

où ρ° est la densité de diffusion de la solution, et v_j le volume de l'atome j . Ce facteur de contraste, $f_j - \rho^\circ v_j$ peut être très différent selon les atomes. Pour un faisceau de photon, il sera très faible pour un hydrogène, mais très important pour un atome lourd (cf. le remplacement multiple isomorphe utilisé en cristallographie). Pour un faisceau de neutrons, il pourra même être nul, ou négatif. Par exemple, la densité de diffusion d'une protéine composée majoritairement de ^1H , de ^{12}C , de ^{14}N est d'environ $2 \cdot 10^{-6} \text{ \AA}^{-2}$. Un mélange d'environ 41% D_2O / 59% H_2O pour la solution a la même densité de diffusion. Le contraste est alors nul ; la protéine devient invisible pour les neutrons. Cette particularité de la diffusion de neutrons permet d'éteindre, en quelque sorte, le signal d'une molécule, en choisissant judicieusement sa composition isotopique et celle de la solution dans laquelle l'expérience est faite. Dans l'étude de complexe de molécules par exemple, cette approche peut aider à visualiser les sous-unités *in situ*.

Le grand flux des sources de rayons X permet d'obtenir des courbes de diffusion avec une très bonne statistique. Les neutrons, bien que les sources aient un flux bien moindre que les rayons X, sont très utiles du fait des possibilités de variation de contraste. L'approche combinée de photons et de neutrons permet donc d'augmenter la précision sur la structure tridimensionnelle des molécules.

III. Les spectroscopies

Les méthodes qui ont été citées précédemment reposent sur les propriétés de diffusion des objets. Les limites de ces méthodes ont été abordées. Une autre manière d'étudier le monde sub-microscopique est d'utiliser les propriétés quantiques des molécules. Une molécule a plusieurs états énergétiques qui sont discrets et quantifiés. Une molécule passe de son état stable à un état excité en absorbant de l'énergie et retourne à son état stable en réémettant de l'énergie. Le domaine qui étudie et utilise ces propriétés d'absorption et d'émission d'énergie est le champ des spectroscopies. Ces méthodes permettent de recueillir des informations, qui peuvent mener à l'élaboration de modèles, et même de structures tridimensionnelles des molécules.

Dans certains domaines, il est plus facile de parler en termes de fréquence plutôt que de longueur d'onde, bien que ces deux notions soient liées par la simple formule de vitesse $c = \lambda \nu$, où c est la vitesse de l'onde, λ sa longueur d'onde, et ν sa fréquence. Il est parfois plus courant de parler de la fréquence des photons car elle est directement proportionnelle à l'énergie qu'ils transportent : $E = h\nu$, avec h la constante de Planck.

On peut séparer les spectroscopies en deux types : celles dites d'absorbance, et celles dites d'émission. L'idée de base est commune et repose sur la mesure d'une absorbance, ou d'une émission, d'énergie en fonction d'un spectre de fréquence ou de longueur d'onde, d'où le nom de spectroscopie.

III.1. La spectroscopie optique

Dans les spectroscopies d'absorbance, l'objet étudié est éclairé, et l'énergie qui a été absorbée est mesurée. En spectroscopie optique, on utilise la gamme de longueur d'onde de l'ultraviolet à l'infrarouge. Quand une molécule est éclairée avec une onde de fréquence UV visible, l'onde électromagnétique va interagir avec les électrons des liaisons chimiques. Ces électrons sont excités par l'énergie de l'onde, d'un niveau orbital électronique fondamental à un niveau plus énergétique. Ces transitions énergétiques sont attribuées à un type de liaisons chimiques et permettent de déduire des informations sur la molécule éclairée.

Par exemple, les électrons des cycles aromatiques des molécules organiques sont excités par des ondes de longueur d'onde entre 230 et 300 nm. On utilise cette absorbance des cycles aromatiques pour mesurer des concentrations de protéines, grâce à la loi de Beer-Lambert. L'intensité émergente I d'un faisceau d'intensité incidente I_0 qui traverse un

échantillon de protéine de largeur l est reliée à la concentration de la protéine c , par le coefficient d'extinction molaire $\varepsilon(\lambda)$.

Eq. 19 :
$$A = \log \frac{I_0}{I} = \varepsilon(\lambda)lc$$

On connaît le facteur d'absorbance des trois types d'acides aminés qui possèdent des cycles aromatiques, le tryptophane, la tyrosine et la phénylalanine en fonction de la longueur d'onde. En connaissant la composition de la protéine, on peut alors calculer le facteur d'absorbance, $\varepsilon(\lambda)$ de la protéine étudiée. Ainsi, la concentration d'un échantillon de cette protéine peut être calculée à partir de son absorbance et de ce facteur $\varepsilon(\lambda)$.

Cet exemple est loin de représenter toutes les applications possibles de spectroscopie UV-visible qui sont très larges. La gamme des longueurs d'onde infrarouge correspond aux niveaux vibrationnelles des liaisons chimiques, cette spectroscopie étudie donc plutôt la dynamique des molécules.

III.2. La résonance magnétique nucléaire

Dans les spectroscopies d'émission, l'objet est excité, préalablement à la mesure. On enregistre l'émission d'énergie lors du retour au niveau fondamental de l'objet.

III.2.1. Les fréquences de résonance

La spectroscopie par Résonance Magnétique Nucléaire (RMN) utilise les transitions entre niveaux énergétiques des spins des noyaux. Le spin d'un noyau peut être vu comme un petit aimant. Tous les atomes ne possèdent pas de spin nucléaire. Le spin nucléaire du carbone ^{12}C et de l'oxygène ^{16}O est nul. Par contre, l'hydrogène ^1H , le carbone ^{13}C , l'azote ^{15}N possèdent un spin $\frac{1}{2}$. Une différence majeure avec la spectroscopie optique est que les niveaux énergétiques magnétiques ne sont pas distincts sans l'application d'un champ magnétique. C'est l'application d'un champ magnétique \vec{B}_0 qui induit l'existence de deux états pour les spins nucléaires : un état où une composante du spin s'oriente parallèlement au champ \vec{B}_0 , et l'autre état, où cette composante s'oriente antiparallèlement. Les spins nucléaires sont répartis entre ces niveaux d'énergie, $E_1 = -\vec{M} \cdot \vec{B}_0$ et $E_2 = \vec{M} \cdot \vec{B}_0$ respectivement, où \vec{M} est le moment magnétique du spin nucléaire. L'échantillon magnétisé peut alors être excité par une

onde si la fréquence correspond à l'énergie de la transition. On parle alors de fréquence de résonance. Les spins excités retournent à leur état d'équilibre en émettant une onde. Ces ondes sont enregistrées, et leur fréquence extraite par transformée de Fourier.

Les spins tournent autour du champ magnétique, dans un mouvement de précession, avec une vitesse de rotation liée à B par l'équation de Larmor

$$\text{Eq. 20 :} \quad \omega = 2\pi\nu = \gamma B$$

où γ est le rapport gyromagnétique, caractéristique d'un noyau. C'est le phénomène de précession de Larmor. Pour être excités, les spins sont basculés dans le plan transverse à \vec{B}_0 par l'application d'un champ transitoire. Ce champ est induit par une impulsion d'onde électromagnétique de fréquence égale à celle de résonance du spin qu'on cherche à basculer. Les spins retournent à leur position d'équilibre en tournant à leur pulsation de résonance.

$$\text{Eq. 21 :} \quad \omega_{res} = 2\pi\nu_{res}$$

On mesure l'aimantation dans le plan transverse, qui est donc oscillante de pulsation ω_{res} , et décroissante.

Cette pulsation n'est pas exactement la même pour tous les spins. En effet la fréquence de résonance d'un spin est proportionnelle au champ magnétique qu'il subit, par le rapport gyromagnétique, γ , caractéristique d'un type de noyau. Ainsi, les protons ^1H , les azotes ^{15}N et le carbone ^{13}C ne résonnent pas dans la même gamme de fréquence, car leur coefficient gyromagnétique est différent. D'autre part, le champ magnétique subit n'est pas équivalent en tout point d'une molécule. Des variations sont induites par l'environnement local du spin. Le spin d'un atome d'hydrogène impliqué dans une liaison hydrogène avec un atome d'oxygène ne subit pas le même champ magnétique qu'un spin au cœur d'une ligne de champ d'un cycle aromatique. Ainsi leurs fréquences de résonance sont légèrement différentes. Ces variations sont de l'ordre de quelques hertz et sont exprimées par rapport à une fréquence de référence ν_{ref} , et normalisées par la fréquence du spectromètre ν_0 , pour obtenir un nombre sans dimensions δ , qui est alors indépendant du spectromètre utilisé.

$$\text{Eq. 22 :} \quad \delta = 10^6 \frac{(\nu_{res} - \nu_{ref})}{\nu_0}$$

δ est appelé le déplacement chimique. La transformée de Fourier du signal enregistré juste après l'impulsion radiofréquence d'excitation donne la composition spectrale du système de spin étudié.

Le spectre des fréquences de résonance des spins des atomes d'hydrogène d'une molécule lui est caractéristique. Il contient les pics de fréquences de résonance de chaque atome d'hydrogène de la molécule. L'attribution de chaque pic à un atome d'hydrogène n'est possible que pour de petites molécules de quelques atomes. Au-delà, le spectre est trop dense.

III.2.2. Les spectres multidimensionnels (HSQC)

Pour réduire le nombre de recouvrement de spins, une méthode consiste à coupler des spins de noyaux d'atomes différents, en se basant sur un couple de fréquence plutôt que sur une fréquence unique. Ce sont les spectres multidimensionnels (Aue *et al.*, 1976).

Les spins nucléaires s'influencent mutuellement, notamment via les liaisons covalentes entre les atomes, on parle de couplage scalaire entre les spins. Ce couplage scalaire est très utile car il permet de coupler plusieurs spins entre eux. En appliquant des impulsions de radiofréquences, les spins peuvent être manipulés. Une succession judicieuse de ces impulsions, appelée séquence INEPT, permet de transférer l'aimantation du spin ^1H au spin du noyau ^{15}N (ou ^{13}C) qui lui est relié covalamment. La fréquence de résonance du spin de ^{15}N (ou ^{13}C) est éditée sur l'aimantation qui est ensuite re-transférée sur le spin du ^1H via le couplage scalaire, par une séquence symétrique à INEPT. Ainsi, la fréquence de résonance d'un spin de ^{15}N ou de ^{13}C , est mesurée indirectement, par le spin d'un ^1H . De cette manière, les deux fréquences de résonance sont corrélées par un même pic. Le pic se trouve dans un plan où ses coordonnées sont les deux fréquences de résonances, celle de l'hydrogène et celle de ^{15}N (ou celle de ^{13}C). Le spectre une dimension ($1\text{D}^1\text{H}$) est alors la projection sur une dimension du spectre deux dimensions [^1H ; ^{15}N (^{13}C)].

Dans le cas d'un spectre [^1H ; ^{15}N], seules les fréquences de résonance des spins des atomes d'hydrogène lié covalamment à un atome de ^{15}N sont éditées. Pour une protéine, on obtient alors autant de pics que d'acides aminés, excepté pour les prolines qui ne possèdent pas de proton amide, plus un pic par hydrogènes amines des chaînes latérales (arginine, glutamine, lysine et asparagine). C'est alors une sorte de signature de la protéine en solution, très sensible à sa conformation. Il est possible à la lecture de ce spectre d'établir si une protéine est structurée ou dans un état dépliée. Ce spectre est dit HSQC.

Comme ce spectre est caractéristique d'une molécule dans un état, on peut aussi examiner des changements de la molécule en fonction de différents paramètres (pH, force ionique...) ou tester des interactions avec des ligands ou d'autres molécules. Ce type de spectre ne suffit pas pour résoudre la structure d'une molécule, mais il apporte des informations qui peuvent aider à l'élaboration d'un modèle tridimensionnel d'un système étudié.

Il est d'autre part possible de résoudre la structure d'une molécule par RMN, à partir d'autres types de spectres. La stratégie ne sera pas développée, mais il est important de savoir que cela est possible pour des molécules de petites masses molaires, jusqu'à 50 kDa environ, ou jusqu'à la centaine de kDa dans des cas particuliers (Tugarinov *et al.*, 2006). La limitation de taille est due au phénomène de relaxation.

III.2.3. La relaxation

Lorsqu'un spin retourne à sa position d'équilibre, le signal émis par le spin, mesuré dans le plan transverse, est alors décroissant. Le temps de décroissance du signal est appelé temps de relaxation transverse, τ_2 .

La résolution des spectres mesurés dépend de ce temps de relaxation. En effet, la largeur des pics de résonance est reliée à ce paramètre. Plus ce temps est court, plus le pic est large. Comme d'autre part, l'aire du pic est conservée (puisque'il représente la somme des spins résonnant à cette fréquence), plus le pic est large, moins il est intense. En dessous d'une intensité minimale, le pic ne peut plus être différencié du bruit (Figure 16).

D'autre part, ce temps de relaxation est lié aux vitesses de réorientation angulaire des molécules par rapport au champ magnétique appliqué et donc, à leur comportement hydrodynamique.

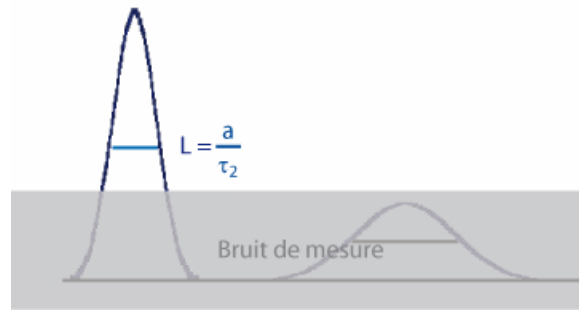


Figure 16 : **Représentation schématique de deux pics de résonance de RMN, de largeur différente.** La largeur à mi-hauteur L est proportionnelle au temps de relaxation transverse τ_2 de la molécule. Un pic trop large n'émerge pas du bruit de mesure symbolisé par un écran gris.

Le temps de diffusion rotationnelle, τ_c , est lié aux dimensions de la protéine :

Eq. 23 :

$$\tau_c = \frac{\eta \cdot V}{k_B \cdot T}$$

où η est la viscosité de la solution, V le volume de la protéine sphérique, k_B , la constante de Boltzmann, T la température.

Plus une molécule est grande, plus les temps de réarrangement en solution sont grands. Plus le temps caractéristique de ces mouvements est grand, plus le temps de relaxation transverse est court (Levitt, 2001). Pour une molécule trop grande, les pics, trop larges, deviennent peu ou pas visibles.

Les spins nucléaires peuvent être vu comme des petits dipôles magnétiques. Ces dipôles induisent des champs magnétiques locaux, qui influencent les spins magnétiques voisins, par ce qu'on appelle alors le couplage dipolaire. Ce couplage dipolaire provoque un déphasage des spins entre eux. Ce déphasage participe à la décroissance du signal émis par l'ensemble des spins dans le plan de mesure. Un moyen de limiter ce déphasage est de diminuer ce couplage dipolaire. Ce couplage est proportionnel au coefficient gyromagnétique des spins. Le rapport gyromagnétique des noyaux de l'isotope 2 de l'hydrogène (le deutérium) étant plus petit que celui de l'isotope 1 (le proton), le couplage dipolaire entre un noyau de deutérium et un proton est plus faible que celui entre deux protons. Ainsi, le remplacement d'une partie des protons d'une molécule par des deutériums permet de diminuer la relaxation des spins des protons et donc d'améliorer la résolution des spectres.

L'utilisation de molécules marquées en isotope ^2H (deutérium) est maintenant courante et a permis d'élargir le champ d'application de la RMN à des molécules plus grandes (Tugarinov *et al.*, 2006).

IV. Le comportement hydrodynamique des molécules

Les molécules biologiques sont dans la nature, la plupart du temps dans un environnement aqueux. Pour les étudier, elles sont généralement manipulées en solution aqueuse. Comme tout objet immergé, elles ont alors un comportement en solution qui obéit aux lois de l'hydrodynamique. L'étude du comportement en solution d'une molécule par la mesure des coefficients hydrodynamiques, est un autre moyen d'obtenir des informations structurales.

IV.1. La diffusion de lumière dynamique (DLS)

La DLS est une technique qui permet de mesurer le coefficient de diffusion D_t , d'une molécule. Les molécules en solution se déplacent de manière chaotique. Dans un temps t , la molécule explore en moyenne, une sphère de rayon r , définit par:

Eq. 24 :
$$\langle r^2 \rangle = 6D_t t$$

où D_t est le coefficient de diffusion translationnel. Il est possible de mesurer ce coefficient en suivant les déplacements d'une molécule, par l'intermédiaire de la lumière diffusée par cette molécule. Une molécule éclairée par un faisceau de lumière, diffuse une partie de cette lumière dans toutes les directions. Les rayons diffusés par les molécules interfèrent et le résultat de ces interférences varie lorsque les molécules se déplacent. Cette variation peut être estimée avec le produit d'auto corrélation de l'intensité diffusée :

Eq. 25 :
$$G(\tau) = \langle I(t)I(t + \tau) \rangle.$$

Ce produit rend compte du temps où le système étudié reste comparable à lui-même, c'est-à-dire pendant lequel il a peu évolué. La décroissance de ce produit est exponentielle, simple ou multiple selon la complexité du système étudié. Le temps caractéristique de cette décroissance exponentielle est relié à la vitesse de diffusion des molécules (plus les molécules se déplacent vite, moins le système reste longtemps comparable à lui-même). Cette vitesse de diffusion est reliée au coefficient de diffusion.

Eq. 26 :
$$G(\tau) = 1 + a \exp(-2Dq^2\tau)$$

Eq. 27:

$$q = n \left(\frac{4\pi}{\lambda} \right) \sin \left(\frac{\theta}{2} \right)$$

Le facteur q tient compte des paramètres de l'expérience : n est l'indice de réfraction de la solution, λ la longueur d'onde utilisée, θ l'angle de diffusion utilisé pour la mesure. (Cette équation est à comparer avec l'équation 3, où l'indice de réfraction où l'indice de réfraction n'apparaît pas explicitement parce qu'il est pratiquement égal à 1 pour des rayons X). Ces paramètres connus, on peut alors extraire une valeur du coefficient de diffusion translationnel.

Pour une sphère en solution, la relation d'Einstein-Stokes établit la relation entre ce facteur et le rayon R de la sphère :

Eq. 28 :

$$D = \frac{k_B T}{f} = \frac{k_B T}{6\pi\eta R}$$

où k_B est la constante de Boltzmann, T la température, f le coefficient de friction de la molécule et η , la viscosité de la solution.

Pour une protéine globulaire, cette formule peut être utilisée, mais certaines molécules ne peuvent être modélisées par une sphère. Cette formule ne donne alors que le rayon d'une sphère qui aurait le même coefficient de diffusion. Ce rayon est appelé rayon hydrodynamique ou rayon de Stokes, et il est caractéristique d'une molécule.

Ainsi, cette méthode permet de calculer une dimension caractéristique d'une molécule, à partir de son comportement hydrodynamique.

IV.2. L'ultracentrifugation analytique (AUC)

Une autre approche de l'étude du comportement hydrodynamique d'une molécule est de mesurer sa vitesse de sédimentation. Lorsqu'une molécule est soumise à une force gravitationnelle ou centrifuge, la vitesse à laquelle elle sédimente dépend de sa résistance au mouvement, c'est-à-dire de son coefficient de friction. L'ultracentrifugation analytique permet de mesurer ce coefficient. La force gravitationnelle qui s'applique sur les molécules est négligeable devant l'agitation thermique, elle n'est donc pas suffisante pour les faire

sédimenter. On applique donc une force centrifuge F_c , de l'ordre de 100 000 fois la gravité terrestre. Le champ de force qui s'applique à la molécule étudiée est donc :

$$Eq. 29 : \quad F_c + F_a + F_f = m\omega^2 r - m_0\omega^2 r - fv$$

où m est la masse de la molécule, r sa distance au centre de rotation, f son coefficient de friction, et v sa vitesse. F_a est la force d'Archimède, F_f la force de friction, et ω la vitesse de rotation.

La vitesse de sédimentation devient constante lorsque ce champ de force s'annule :

$$Eq. 30 : \quad m\omega^2 r - m_0\omega^2 r - fv = 0$$

$$Eq. 31 : \quad (m - m_0)\omega^2 r = m \left(1 - \frac{\rho_0}{\rho} \right) \omega^2 r = m(1 - \bar{v}\rho_0)\omega^2 r = fv$$

où m_0 est la masse du solvant déplacé par la molécule, ρ est la masse volumique de la molécule, ρ_0 celle de la solution, \bar{v} le volume spécifique de la molécule ($\rho = v.\rho_0$). On définit un paramètre, le coefficient de sédimentation s , comme le rapport entre la vitesse de sédimentation et l'accélération subit par la molécule. Ce paramètre est alors indépendant de la force centrifuge appliquée :

$$Eq. 32 : \quad s = \frac{v}{\omega^2 r} = \frac{M(1 - \bar{v}\rho)}{N_A f}$$

où M est la masse molaire de la molécule, et N_A la constante d'Avogadro. Cette grandeur a la dimension d'un temps et s'exprime en Svedberg, qui correspond à 10^{-13} secondes (Serdyuck *et al.*, 2007). Ce coefficient permet de calculer le coefficient de friction de la molécule, en présupposant sa masse.

Ce coefficient de sédimentation est mesuré par le suivi de l'absorbance d'un échantillon de concentration connue. Le front d'absorbance s'éloigne du centre de rotation lorsque les molécules sédimentent, en fonction du temps de rotation. La vitesse de progression du front d'absorbance, en fonction de la vitesse de rotation appliquée, donne le coefficient de sédimentation :

Eq. 33 :

$$s = \frac{1}{\omega^2} \frac{d \ln(r_f)}{dt}$$

où r_f est la distance au centre de rotation du front d'absorbance.

Le coefficient de friction f , pour une sphère, relié à son rayon R par l'équation de Einstein Stokes :

Eq. 34 :

$$f = 6\pi\eta R$$

Ainsi, comme la DLS, cette méthode permet d'obtenir le rayon de la sphère qui aurait le même comportement hydrodynamique. Il est important de noter que ce calcul nécessite une hypothèse de départ sur le degré d'oligomérisation de la molécule en solution : la masse de l'espèce peut être le double de la masse de la molécule dans le cas d'un dimère par exemple.

L'ultracentrifugation analytique est aussi un moyen de constater l'homogénéité ou l'hétérogénéité d'un échantillon. En effet, l'analyse des fronts de migration donne la distribution de quantité de matière en fonction du coefficient de sédimentation. Plusieurs espèces dans un échantillon peuvent être distinguées si elles n'ont pas le même coefficient de sédimentation.

Les méthodes citées ci-dessus, DLS et AUC, donnent accès, entre autre, à une dimension caractéristique des molécules, grâce aux mesures de leurs coefficients hydrodynamiques. Cette dimension, le rayon de Stokes, n'est pas, à proprement parler, une dimension réelle, car elle représente la dimension de la sphère de comportement analogue. Elle est très utilisée, notamment pour comparer les molécules entre elles, et comme premier aperçu de la taille d'une molécule, mais ne doit pas être sur interprétée.

V. L'analyse d'interaction de biomolécules (BIA) par résonance plasmonique de surface (SPR)

Un autre moyen d'étudier les molécules (et de tenter de les représenter dans l'espace) est de les fixer sur une surface et de suivre les variations des propriétés de cette surface. On peut par exemple, utiliser les propriétés de réflexion d'onde électromagnétiques grâce au phénomène de résonance plasmonique de surface.

Quand on éclaire une surface séparant deux milieux transparents d'indice de réfraction différents, une partie du rayon incident est réfléchi, une autre est réfractée. Cette partition est régie par les lois de Snell-Descartes ; notamment l'angle de réfraction θ_2 , par la formule $n_1 \sin(\theta_1) = n_2 \sin(\theta_2)$, où n_1 et n_2 sont les indices de réfraction des milieux, θ_1 l'angle d'incidence du rayon. Si l'indice du milieu de l'onde incidente est supérieur à celui de l'autre milieu ($n_1 > n_2$), alors au-delà d'un certain angle d'incidence $\theta_1 = \text{Arcsin}\left(\frac{n_2}{n_1}\right)$, il y a réflexion totale. Seule une fraction de l'onde électromagnétique pénètre dans le milieu de faible indice de quelques longueurs d'onde seulement. Cette onde est appelée onde évanescente car l'amplitude de cette onde décroît avec la distance parcourue.

Si une fine couche de métal est déposée sur la surface, l'onde évanescente d'une réflexion totale fait entrer en résonance les électrons du métal et cette résonance absorbe l'énergie de l'onde. Paradoxalement, l'intensité de ce rayon totalement réfléchi est fortement diminuée. On parle de réflexion frustrée (Figure 17). Ce phénomène d'extinction se produit à un angle d'incidence très précis ; cet angle dépend des constantes diélectriques du métal et du milieu éclairé (verre du système optique), et de l'indice de réfraction du milieu du côté non éclairé du métal (Serdyuck *et al.*, 2007). Ce phénomène de résonance plasmonique de surface permet donc de sonder un milieu via son indice de réfraction de manière non invasive.

L'indice de réfraction d'un milieu dépend, des éléments qui s'y trouvent. Une variation de concentration de certaines molécules va induire une variation d'indice de réfraction. On peut donc mesurer grâce à ce phénomène des variations de concentrations de molécules dans un milieu, dans une zone proche de la surface de mesure. C'est pourquoi il est utilisé en biologie, pour mesurer et analyser des interactions entre biomolécules, fixées sur la surface et circulantes.

Dans les instruments d'analyse d'interaction de biomolécules (BIA), des molécules (ligands) sont immobilisées à la surface du capteur. L'immobilisation peut être faite par affinité avec la surface du bio capteur, via une étiquette d'acides aminés par exemple. D'autres molécules (analytes) sont entraînées par un système micro fluide, sur la surface du bio capteur, par un flux liquide et continu. Les interactions sont détectées grâce à la mesure de l'angle d'extinction. Le capteur est éclairé avec une gamme de rayon incident inférieur à l'angle critique, tous sont totalement réfléchis. Un détecteur mesure l'intensité réfléchi en fonction de l'angle de réflexion. À l'angle où le phénomène de résonance plasmonique de surface se produit, une baisse nette et aiguë de l'intensité est mesurée.

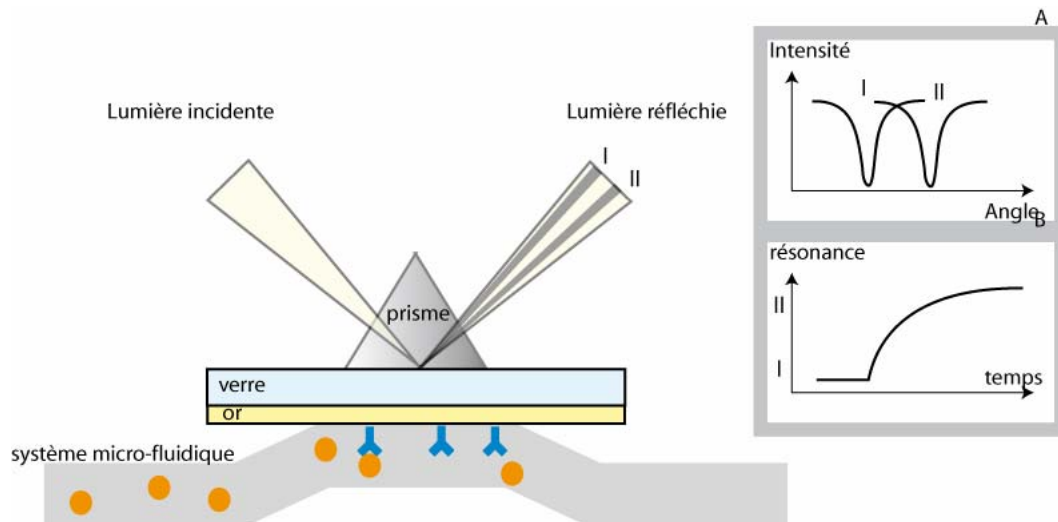


Figure 17 : Principe de l'analyse d'interaction de biomolécules par résonance plasmonique de surface. Le système de micro fluidique est posé sur le senseur composé d'une feuille d'or et d'une couche de verre. Un prisme permet d'éclairer cette surface avec une gamme de rayons incidents à différents angles d'incidence, les rayons réfléchis sont détectés excepté le rayon frustré (I et II). A : la baisse d'intensité à l'angle de la résonance plasmonique de surface est suivie et représentée en fonction de l'angle. B : cette variation est ensuite représentée en fonction du temps, pour pouvoir suivre l'évolution du système en fonction d'injection de molécules (analytes) sur la surface sondée, via le canal d'injection.

Les variations de cet angle en fonction du temps permettent de suivre les événements qui se produisent à la surface du senseur. Une interaction entre les molécules circulantes (analytes) et les molécules immobilisées (ligands), provoque un changement de masse à la surface du bio senseur, qui se traduit par une variation d'indice de réfraction et donc de l'angle d'extinction. L'unité des variations de l'angle de résonance est arbitraire. La valeur absolue de l'angle contient peu d'informations par rapport à ses variations. De ces variations peuvent être extraites les constantes de vitesse d'association et de dissociation des interactions, ainsi que les constantes d'équilibre. Ces variations peuvent être très fines, les instruments les plus récents détectent de variations de masse de 200 Da, et permettent de mesurer des interactions très fortes (K_D jusqu'à quelques nM) et très faibles (K_D de quelques mM).

Cette méthode permet d'étudier et de tester les interactions entre biomolécules en temps réel. Elle permet aussi, de manière indirecte, de cartographier une interaction moléculaire, dans une étude comparée d'une molécule et de mutants dérivés (Beattie *et al.*, 2008). Cette cartographie peut donner, dans une certaine mesure, des informations sur une zone d'interaction dans un complexe. On peut ainsi, construire ou améliorer un modèle tridimensionnel des molécules étudiées.

C. Présentation du modèle utilisé et objectifs du travail

Peu de choses sont connues sur les protéines impliquées dans la division DivIB, DivIC et FtsL. L'objectif de ce travail était d'apporter des informations structurales sur ces trois protéines.

Pour cela, nous avons utilisé le système modèle mis au point par Marjolaine Noirclerc-Savoie et André Zapun (Noirclerc-Savoie *et al.*, 2005).

I. Présentation du modèle

Le système modèle utilisé a été construit à partir des protéines de *S. pneumoniae*, et est exprimé de manière recombinante chez *E. coli* : le domaine extracellulaire de DivIB, le domaine extracellulaire de DivIC exprimé en fusion avec un peptide artificiel, le domaine extracellulaire de FtsL exprimé en fusion avec un peptide artificiel.

Les peptides artificiels e5 et k5 sont constitués de 5 répétitions du motif e, EVSALEK et k, KVSALKE. Ces peptides sont connus pour former une structure appelée « coiled-coil », via l'enfouissement des acides aminés hydrophobes L et V dans une structure en double hélice (Figure 18). La formation du complexe hétérodimérique est assurée par leur complémentarité de charge, avec une constante de dissociation de l'ordre du nanomolaire (Chao *et al.*, 1996).

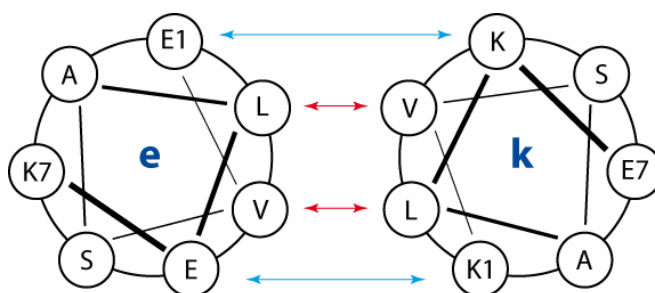
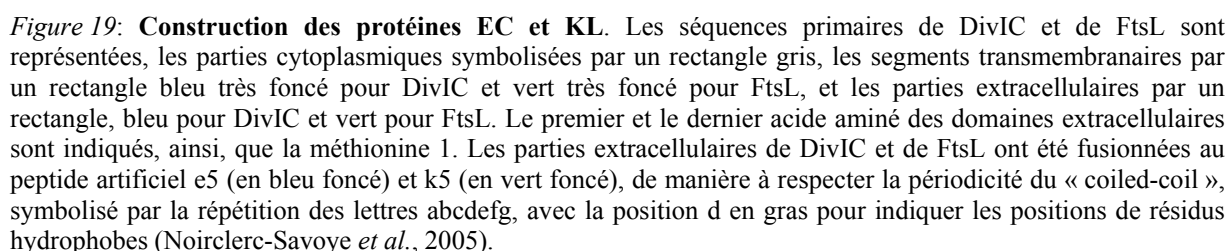


Figure 18 : Représentation schématique du « coiled-coil » formé par le motif e et le motif k. La direction de propagation de la chaîne peptidique, du N-ter au C-ter, est symbolisée par des traits d'épaisseur croissante à l'intérieur du cercle symbolisant l'hélice du résidu indexé 1 au résidu indexé 7. Les flèches rouges représentent les interactions hydrophobes, les bleues, les interactions électrostatiques (Chao *et al.*, 1996).



La validité de ce modèle pour le dimère FtsL/DivIC a été testée par l'interaction avec le domaine extracellulaire de DivIB. Plusieurs combinaisons de fusion ont été testées. Le fait que seul le dimère contenant les domaines extracellulaires de DivIC et de FtsL ait interagi avec le domaine extracellulaire de DivIB, valide ce modèle (Noirclerc-Savoye *et al.*, 2005).

Le système modèle est donc constitué de trois protéines recombinantes, construites à partir des protéines DivIB, DivIC et FtsL de *S. pneumoniae* : IB, EC et KL (Figure 20).

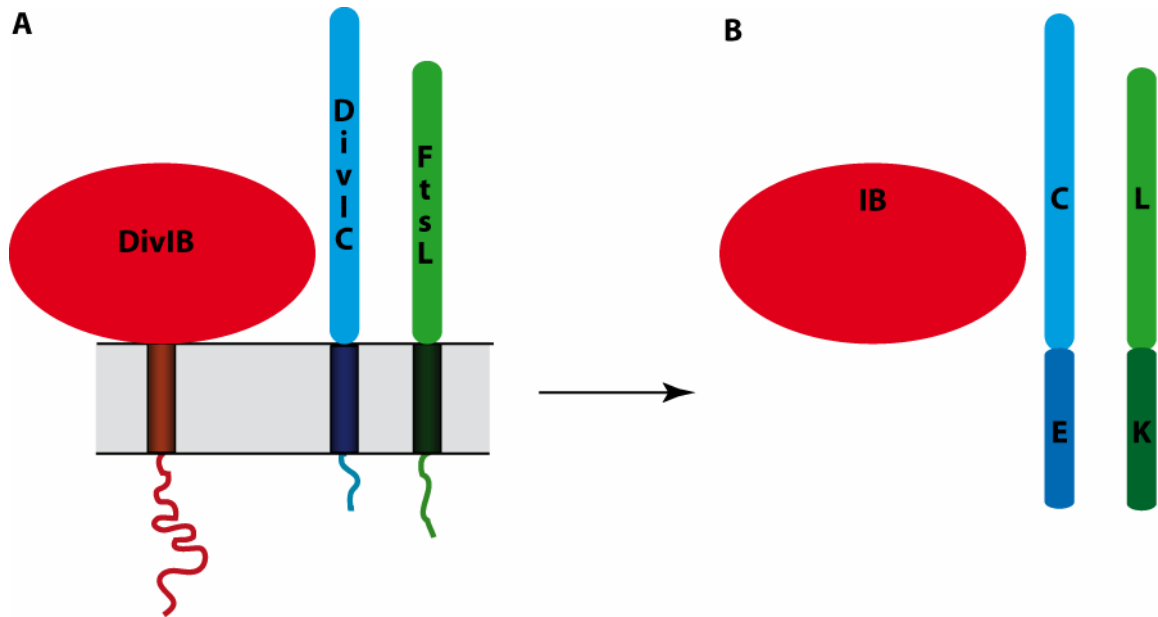


Figure 20 : Schéma représentant le système de protéines utilisées. Les protéines recombinantes IB (en rouge), EC (en bleu) et KL (en vert) (B), ont été construites à partir des protéines DivIB (en rouge), DivIC (en bleu) et FtsL (en vert) de *S. pneumoniae* (A).

II. Objectif du travail

Le but de cette étude était de pouvoir déduire des informations sur l'association des 3 protéines DivIB, DivIC et FtsL dans la bactérie au moment de sa division.

Nous avons mené une étude structurale sur ce modèle. Les objectifs du travail étaient d'apporter des informations structurales sur :

- 1) Le complexe contraint de DivIC et de FtsL, EC/KL (la structure de ce complexe, et l'arrangement des 2 protéines dans le complexe).
- 2) Le domaine extracellulaire de DivIB, IB.
- 3) L'interaction entre le complexe contraint des domaines extracellulaires de DivIC et de FtsL, EC/KL et le domaine extracellulaire de DivIB, IB. Le but était de déterminer l'interface de cette interaction.
- 4) Le complexe formé par les domaines extracellulaires de DivIB, IB, et le complexe contraint des domaines extracellulaires de DivIC et de FtsL, EC/KL. L'objectif était de trouver la structure de ce complexe et l'arrangement des 3 protéines dans ce complexe.

Chapitre 1. Le complexe contraint EC/KL des domaines extracellulaires de DivIC et de FtsL de *S. pneumoniae*

Nous avons mené une étude structurale sur le complexe contraint des parties extracellulaires des protéines DivIC et FtsL de *S. pneumoniae*, EC/KL (cf. Introduction/C.). L'objectif de cette étude était de déterminer la structure tridimensionnelle de ce complexe, à la meilleure résolution possible.

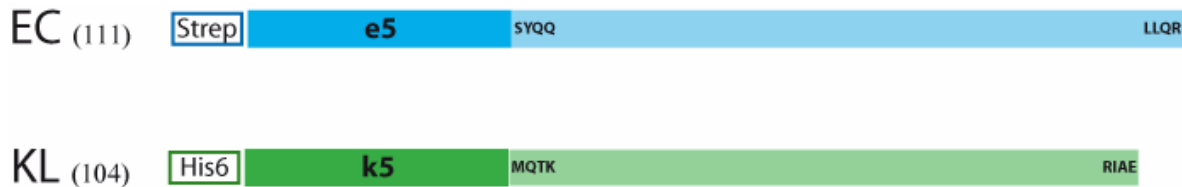


Figure 21 : Schéma des protéines élaborées à partir des domaines extracellulaires de DivIC et FtsL de *S. pneumoniae*. La protéine EC est la fusion du peptide e5 (bleu foncé) avec le domaine extracellulaire de DivIC, noté C (bleu clair). La protéine KL est la fusion du peptide k5 (vert foncé) avec le domaine extracellulaire de FtsL, noté L (vert clair). En face des constructions, est indiqué le nom utilisé, et entre parenthèses, le nombre de résidus de ces constructions. Les premiers et les derniers acides aminés des domaines extracellulaires de DivIC et de FtsL sont indiqués, ainsi que les étiquettes ajoutées pour permettre les purifications par chromatographie d'affinité, Strep pour EC et hexahistidine pour KL.

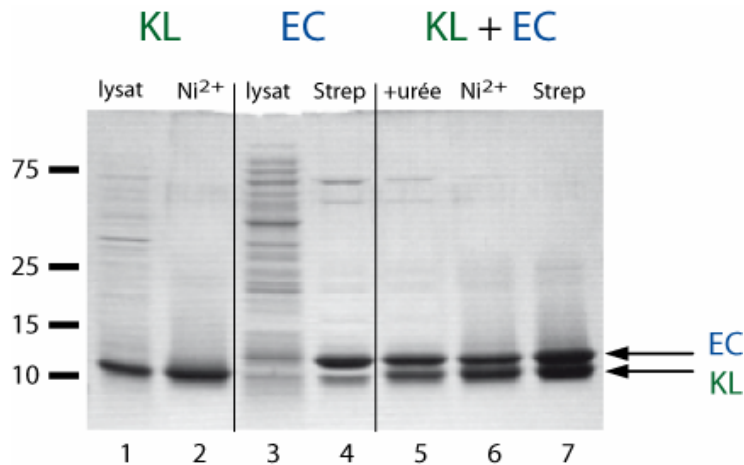


Figure 22 : Co-purification du complexe EC/KL. SDS-PAGE 15% coloré au bleu de Coomassie des fractions purifiées de KL (piste 2) et EC (piste 4), à partir des lysats bactériens d'*E. coli* (pistes 1 et 3), avant mélange en présence d'urée (piste 5). On remarque la présence d'une bande contaminante vers 10 kDa, dans la fraction purifiée de EC. Le mélange a ensuite été purifié sur résine Ni-NTA, l'élution de la purification sur Nickel (piste 6) a ensuite été purifiée sur Strepactin, l'élution de la purification sur Streptactin (piste 7) montre bien la présence de EC et de KL, indiquée par les flèches. Les masses molaires sont indiquées à gauche du gel en kDa

I. Co-purification du complexe EC/KL

Les protéines utilisées sont EC et KL. La protéine EC, est une fusion entre le domaine extracellulaire de DivIC de *S. pneumoniae*, de la sérine 55 à l'arginine 122 (67 acides aminés), avec un peptide appelé e5 (Chao *et al.*, 1996), de 32 acides aminés. La protéine KL est une fusion entre le domaine extracellulaire de FtsL de *S. pneumoniae*, de la méthionine 44 au glutamate 105 (61 acides aminés), avec un peptide appelé k5 (Chao *et al.*, 1996), capable d'interagir avec le peptide e5, de 32 acides aminés (cf. Introduction/C.) (Figure 19). Pour confirmer la formation d'un complexe entre EC et KL, les protéines ont été exprimées séparément dans *E. coli* en fusion avec des étiquettes de purification différentes. La protéine KL a été encodée en fusion N-terminale avec une étiquette hexahistidine, permettant une purification par chromatographie d'affinité au nickel, la protéine EC avec une étiquette Strep permettant une purification par chromatographie d'affinité à la Streptactin® (Figure 21). Ainsi, la protéine KL, avec son étiquette His6, est constituée de 104 acides aminés et sa masse molaire prédite est de 11 718 Da. La protéine EC, avec son étiquette Strep, est constituée de 111 acides aminés et sa masse molaire prédite est de 12 932 Da.

Chacune des protéines a été purifiée indépendamment par chromatographie d'affinité en utilisant une résine appropriée à son étiquette de purification (Figure 22). Les protéines ont ensuite été mélangées en présence d'un agent chaotrope (de l'urée à 8M), pour permettre leur dénaturation complète. L'abaissement de la concentration d'urée par dialyse a ensuite permis la co-structuration des protéines, par leur renaturation en présence l'une de l'autre. Le mélange des deux protéines a été purifié par chromatographie d'affinité sur une résine Ni-NTA®. L'éluat de la chromatographie contient non seulement la protéine KL, mais aussi la protéine EC, indiquant l'interaction entre ces deux protéines (Figure 22). Cette interaction est ensuite confirmée par la co-élution de la protéine KL avec la protéine EC, d'une chromatographie d'affinité à la Streptactin® qui fixe l'étiquette Strep, qui n'est portée que par la protéine EC (Figure 22).

Le rapport d'intensité des bandes EC et KL après double co-purification, indique que la stoechiométrie du complexe EC/KL est un rapport 1/1. La pureté du complexe est estimée supérieure à 95%.

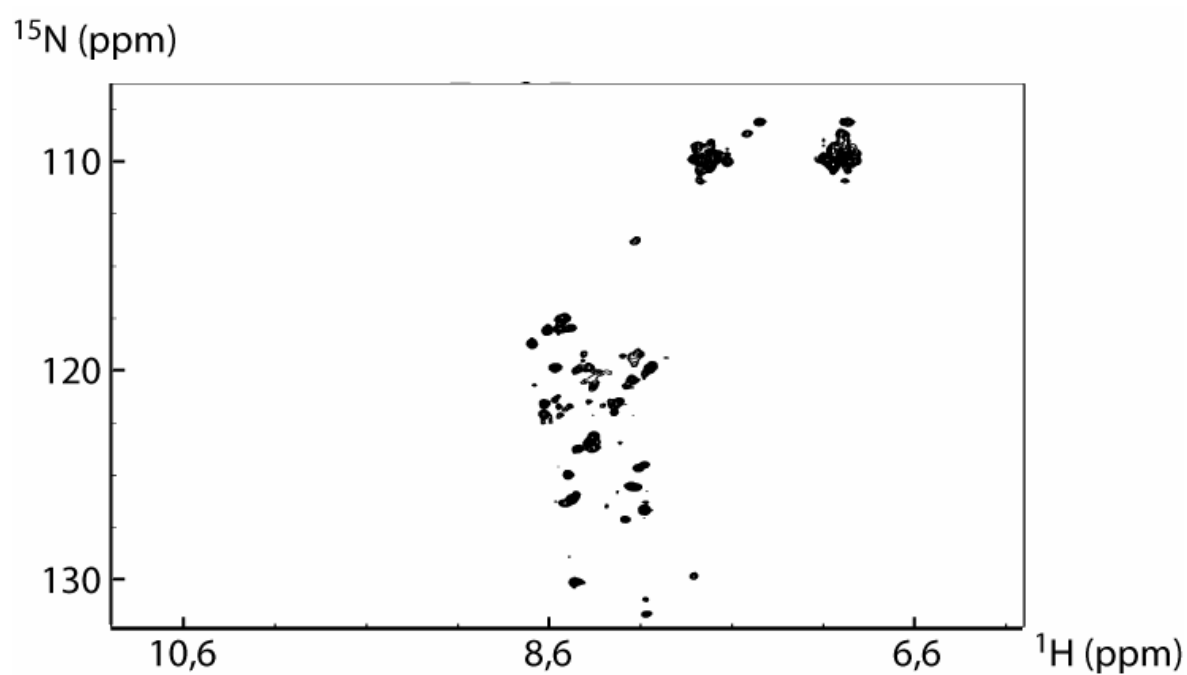


Figure 23 : Spectre HSQC [^{15}N ; ^1H] de la protéine KL dans le complexe EC/KL. L'échantillon EC(^{14}N)/KL(^{15}N) était à une concentration de 5,8 mg/mL dans une solution tamponnée : 20 mM Tris pH 8, 150 mM NaCl, 90% H_2O /10% D_2O .

II. Analyse structurale et biochimique

II.1. Première approche par RMN

Nous avons commencé l'étude structurale de ce complexe EC/KL par une approche par RMN. Le spectre 1D ^1H des fréquences de résonance des protons d'un échantillon de EC/KL montrant quelques indices de protéines structurées, l'étude a été poursuivie par l'enregistrement d'un spectre HSQC [^{15}N ; ^1H]. Pour enregistrer ce spectre, il faut marquer au moins une des deux protéines avec de l'azote 15. La protéine KL s'exprimant et se purifiant en plus grande quantité que EC, c'est cette protéine que nous avons choisi d'exprimer en milieu minimum enrichi en ^{15}N . Un échantillon EC(^{14}N)/KL(^{15}N) a donc été préparé.

Peu de pics se distinguent du bruit, environ 35 (Figure 23), sans compter les pics des azotes des chaînes latérales. La protéine KL, marquée ^{15}N qui est donc la seule sondée, compte 103 acides aminés, dont seulement 35 acides aminés environ sont distingués sur ce spectre.

La structure de KL est prédite pour former un « coiled-coil » avec la protéine EC. Un « coiled-coil » s'étendant sur 70-80 acides aminés serait fortement anisotrope, et le temps caractéristique des mouvements de la molécule autour le petit axe de ce « coiled-coil » pourrait être suffisamment grand pour, via un temps de relaxation court, rendre indétectables les pics de résonance des protons amides. Ces pics élargis sont indistinguables du bruit de fond de la mesure. Pour valider cette hypothèse, nous avons cherché à déterminer les dimensions caractéristiques du complexe EC/KL, par une combinaison de méthodes biophysiques.

II.2. Insertion d'un résidu aromatique dans la protéine KL

Pour ces expériences, nous avons inséré par mutagenèse dirigée un codon pour un tryptophane entre l'étiquette de purification hexahistidine et le peptide K, dans le gène codant pour la protéine KL, afin de pouvoir détecter cette protéine par absorbance à 280 nm. En effet, sans acide aminé aromatique dans sa séquence primaire, la protéine KL n'absorbe pas à 280 nm, son élution de différentes chromatographies ne pouvait donc pas être détectée, ni sa concentration, estimée. Après insertion de cet acide aminé, la protéine KL est alors constituée de 105 acides aminés, et sa masse moléculaire est de 11 905 Da et peut désormais être suivie par son absorbance à 280 nm.

	Rayon de Stokes (Å)	Paramètre de polydispersité
EC/KL 1 mg/ml	$43,5 \pm 0,7$ (30 mesures)	18 %
EC/KL 2,25 mg/ml	$44,9 \pm 0,9$ (33 mesures)	22 %
EC/KL 3,5 mg/ml	$42,6 \pm 0,6$ (31 mesures)	22 %

Tableau 2 : Valeurs du rayon de Stokes d'échantillons de complexe EC/KL. Les valeurs de rayon sont la moyenne sur environ 30 mesures. L'erreur de cette valeur est l'écart type de cette moyenne. Le paramètre de polydispersité est la moyenne des erreurs sur les mesures.

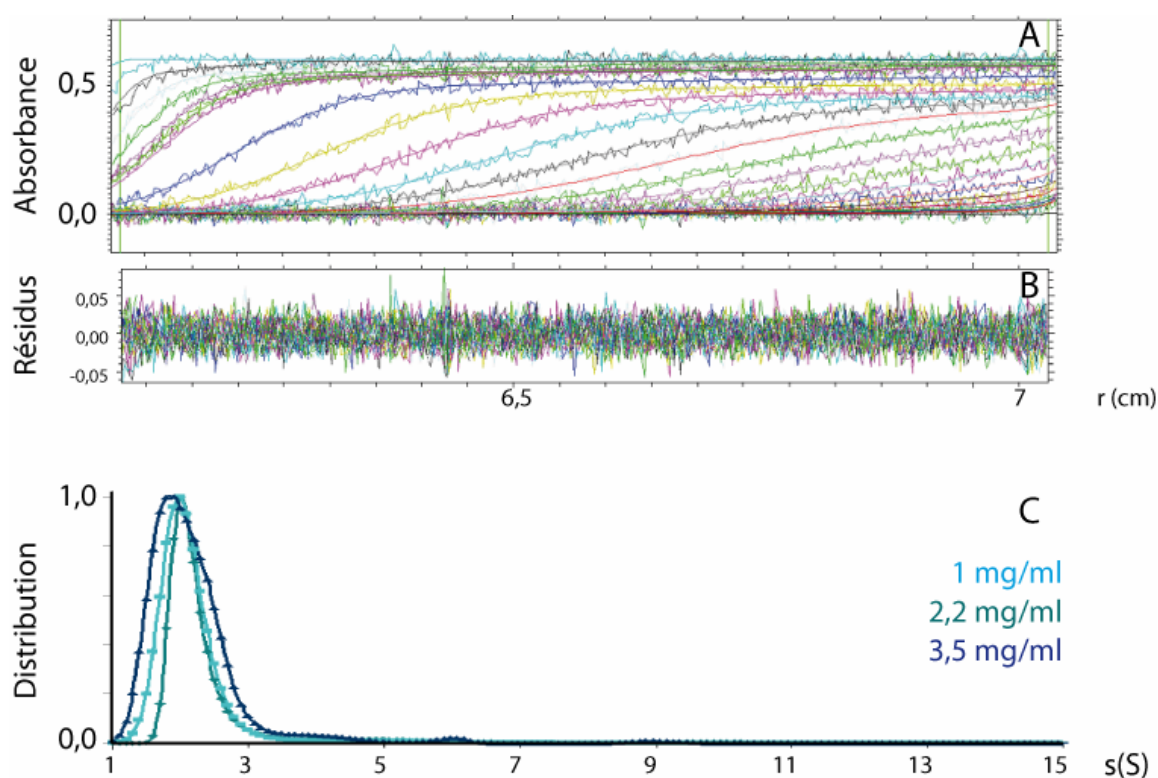


Figure 24 : Ultracentrifugation analytique, vitesses de sédimentation d'échantillons de complexe EC/KL. A : superposition de profils de sédimentation expérimentaux d'un échantillon de EC/KL à 1 mg/ml dans Tris 20 mM pH 8, NaCl 150 mM, à 42 000 rpm à 20°C, et des profils simulés et ajustés par le logiciel Sedfit. B : résidus de l'ajustement. C : distribution en fonction de s (S) de EC/KL à 1 (en bleu clair), 2,2 (en bleu-vert) et 3,5 (en bleu foncé) mg/ml.

II.3. Estimation de la taille du complexe EC/KL (DLS, SEC, AUC)

Tout d'abord, nous avons mesuré le coefficient de diffusion translationnel, par diffusion de lumière dynamique (DLS), de plusieurs échantillons de EC/KL de 1, 2,2 et 3,5 mg/ml, pour vérifier une potentielle évolution de la composition des échantillons en fonction de la concentration. Une trentaine de mesures ont été réalisées pour chaque échantillon. Les valeurs du rayon de Stokes sont reportées dans le tableau 2.

L'erreur de mesure du rayon de stokes n'excède pas 22 % de la valeur moyenne, indiquant que les échantillons pouvaient être considérés comme monodisperses. Le rayon de Stokes calculé peut donc être attribué au complexe EC/KL. Ce rayon est comparable pour les trois échantillons, indiquant que la composition de échantillons ne semble pas évoluer dans cette gamme de concentration (1 à 3,5 mg/ml). Ce rayon est d'environ 43 Å.

Le rayon de Stokes est le rayon d'une sphère qui aurait le même comportement hydrodynamique que la molécule étudiée. Or, le rayon d'une protéine de 25 kDa, correspondant à un complexe (EC/KL)₁ de masse $M_{EC} + M_{KL}$, parfaitement sphérique (cas non réaliste), serait d'environ 20 Å. On remarque donc que le rayon de Stokes du complexe EC/KL est environ 2 fois plus grand. On peut donc considérer que la forme du complexe est éloignée d'une sphère. Le complexe est donc non compact et/ou non globulaire, ou partiellement non structuré. Il est aussi possible que le complexe soit un oligomère du complexe (EC/KL)₁. Ces possibilités sont également suggérées par le fait qu'un échantillon de EC/KL est élué d'une chromatographie par exclusion de taille (SEC), comme une protéine globulaire d'environ 50 kDa.

Pour compléter cette caractérisation, nous avons aussi réalisé des expériences d'ultracentrifugation analytique. L'analyse de sédimentation de trois échantillons de EC/KL de 1, 2,2 et 3,5 mg/ml à 20°C par le logiciel Sedfit (Schuck, 2000) montre une seule espèce apparente dans chaque échantillon (Figure 24).

Cette espèce moléculaire apparente sédimente avec un coefficient de sédimentation de $2,05 \pm 0,08$ S pour les trois échantillons. On peut donc conclure que, dans cette gamme, la concentration n'a pas d'effet sur l'oligomérisation du complexe EC/KL.

Cette espèce apparente est de grande taille. En effet, l'ajustement des simili-courbes des profils de sédimentation des trois échantillons n'est optimal que lorsque le paramètre du rapport des coefficients de friction réel f et de la protéine idéalement sphérique f_0 , (f/f_0), est laissé libre d'évoluer, et ce paramètre évolue systématiquement vers des valeurs comprises entre 1,6 et 1,8. Le complexe EC/KL est donc de forme assez éloignée d'une sphère, ce qui avait déjà été suggéré par les résultats de DLS. Plus précisément le rapport f/f_0 indique (pour

	Avec f/f_0 extrait de l'AUC	Avec R_S extrait de la DLS	Avec une colonne de SEC calibrée
M_{app}	33 kDa environ	43 kDa environ	50 kDa environ

Tableau 3 : Masse apparente du complexe EC/KL déterminée par différentes méthodes. Les calculs de la masse à partir du rapport des coefficients de friction (Eq. 35) et du rayon de Stokes déterminé par DLS (Eq. 36) ont été réalisés avec $\eta = 1.02$ poise, $\rho = 1.005$ g/cm³ et $v = 0,74$ cm³/g. La détermination de la masse par SEC a été réalisée par calibration de la colonne (cf. chapitre 4/IV.3).

un ellipsoïde rigide) un rapport (grand axe)/(petit axe) de l'ordre de 12 à 15 (Schachman, 1959).

La masse de cette espèce n'a pas pu être déterminée de façon non ambiguë. En effet, les estimations par différentes méthodes n'ont pas donné la même valeur (Tableau 3).

En utilisant le coefficient de sédimentation et le rapport des coefficients de friction réel et de la protéine idéalement sphérique, déterminés par l'ultracentrifugation analytique, on peut déterminer la masse apparente de l'espèce par la formule (cf. Annexe):

$$Eq. 35 : \quad M_{app} = \left(\frac{s \cdot N_A \cdot 6\pi}{1 - \rho \bar{v}} \cdot \frac{f}{f_0} \right)^{\frac{3}{2}} \left(\frac{3\bar{v}}{4\pi N_A} \right)^{\frac{1}{2}}$$

En utilisant le rayon de Stokes déterminé par la diffusion dynamique de lumière et le coefficient de sédimentation déterminé par ultracentrifugation analytique, on peut déterminer la masse apparente de l'espèce par la formule :

$$Eq. 36 : \quad M_{app} = \frac{s \cdot N_A \cdot 6\pi \eta \cdot R_s}{1 - \rho \bar{v}}$$

Dans ces deux équations, s est le coefficient de sédimentation, ρ , la masse volumique du solvant, \bar{v} , le volume spécifique de la protéine, η , la viscosité du solvant, N_A la constante d'Avogadro, f , le coefficient de friction de la protéine, f_0 le coefficient de friction de la protéine idéalement sphérique, R_s , le rayon de Stokes de la protéine.

En utilisant le volume d'élution de l'espèce d'une colonne de chromatographie par exclusion de taille, on peut déterminer la masse apparente de l'espèce par la droite résultant de la calibration. Les résultats des calculs sont réunis dans le tableau 3.

Il est important de remarquer que toutes les méthodes utilisées pour déterminer la masse de ce complexe reposent sur le comportement hydrodynamique des protéines.

Quelque soit la masse moléculaire du complexe, ces résultats indiquent que le complexe EC/KL est de grande taille. Ceci peut expliquer l'absence des pics de résonance de KL du spectre en figure 23 (cf. chapitre1/II.1). Ce complexe EC/KL ne pourrait donc pas être étudié par RMN de manière classique. C'est pourquoi, nous avons décidé de poursuivre cette étude structurale par diffusion aux petits angles. Tout d'abord, nous avons voulu mesurer la

masse molaire du complexe EC/KL par cette technique qui ne repose pas sur le comportement hydrodynamique des protéines. Nous avons également tenté d'obtenir la structure basse résolution de ce complexe. Nous avons choisi d'utiliser la diffusion de neutrons, afin d'utiliser les propriétés de contraste dues au marquage isotopique ^2H , et tenter de localiser les protéines EC et KL à l'intérieur du complexe.

III. Élaboration des constructions génétiques et des protocoles de purification pour les expériences de diffusion aux petits angles

III.1. Marquages isotopiques

Pour utiliser les propriétés de contraste du neutron, nous avons voulu réaliser un marquage isotopique ^2H , ou deutération, des protéines EC et KL. Pour cela, les bactéries utilisées pour produire les protéines ont été cultivées en eau deutérée, $^2\text{H}_2\text{O}$ (D_2O). La croissance des bactéries en D_2O est plus lente, le protocole de production est donc plus long qu'en H_2O et nécessite une grande stabilité de la pression de sélection des bactéries par les antibiotiques, garant de la maintenance des plasmides codant pour les protéines recombinantes. La protéine EC étant codée sur un plasmide induisant une résistance à l'antibiotique kanamycine, le marquage de cette protéine pouvait être réalisé sans modification du système d'expression. Par contre, le plasmide d'expression d'origine de la protéine KL induit une résistance à l'antibiotique ampicilline, insuffisamment stable. En effet, la résistance à cet antibiotique est permise par l'expression d'une bêta-lactamase dont l'action est extracellulaire et provoque la baisse de la concentration totale de l'antibiotique dans le milieu de culture, diminuant la pression de sélection. La résistance encodée par le vecteur d'expression de KL, à l'ampicilline a donc été remplacée par celle à la kanamycine, à l'aide du système EZ-Tn5TM. Ce système permet l'insertion d'une séquence, ici la séquence codant pour une protéine induisant une résistance à la kanamycine, dans le chromosome ou les plasmides des bactéries, à une position aléatoire. Le plasmide d'expression de KL [^RKan ; ^SAmp] (où la séquence de résistance à la kanamycine a été insérée dans la séquence de résistance à l'ampicilline, supprimant cette dernière), a ensuite été isolé par sélection d'une souche d'*E.coli* BL21 DE3 transformée avec ce vecteur poussant en présence de kanamycine et ne poussant pas en présence d'ampicilline.

Les deux protéines EC et KL ont ensuite été produites en bioréacteur, séparément, en milieu deutéré pour réaliser un marquage ^2H (D) des protéines.

La production séparée des protéines a permis de préparer des échantillons de complexe EC/KL avec différentes combinaisons de marquage. Les protéines, marquées ou non, ont été purifiées séparément, puis mélangées, afin de préparer des complexes complémentaires EC(D)/KL(H) et EC(H)/KL(D). Les échantillons partiellement deutérés ont ensuite été transférés dans un tampon 41% D_2O /59% H_2O afin de réaliser l'extinction de contraste de la protéine hydrogénée. En effet, la densité de diffusion d'une solution composée de 41 % de D_2O est équivalente à la densité de diffusion d'une protéine hydrogénée (Jacrot, 1976). Ainsi, une protéine hydrogénée est invisible dans ce type de solution, et seule la diffusion d'une protéine deutérée peut être mesurée.

III.2. Co-expression

Afin de simplifier la purification du complexe EC(H)/KL(H), une construction génétique comportant les gènes codant pour les deux protéines a été réalisée. Le gène codant pour la protéine EC a été inséré dans le plasmide d'expression de KL, permettant ainsi la co-expression des deux protéines. Cette co-expression des protéines EC et KL permet leur association probable dans la bactérie, pouvant accroître leur stabilité réciproque en les protégeant mutuellement de la dégradation. D'autre part, cette co-expression annule la nécessité de purification préalable des protéines avant mélange, et a donc permis d'optimiser le protocole de purification du complexe EC/KL. Le protocole est allégé de deux étapes de chromatographie d'affinité et de plusieurs dialyses. Le rendement est ainsi multiplié par 2 environ (de 10 mg à 20 mg de complexe purifié environ, par litre de culture). Toutefois, cette co-expression ne permet pas le marquage spécifique d'une des deux protéines.



Figure 25 : Estimation de la pureté des échantillons de complexe EC/KL dans les différentes combinaisons de marquage isotopique. SDS-PAGE 15% coloré au bleu de Coomassie de EC(H)/KL(H) en piste 1, EC(D)/KL(H) en piste 2 et EC(H)/KL(D) en piste 3. Les masses molaires sont indiquées à gauche du gel en kDa.

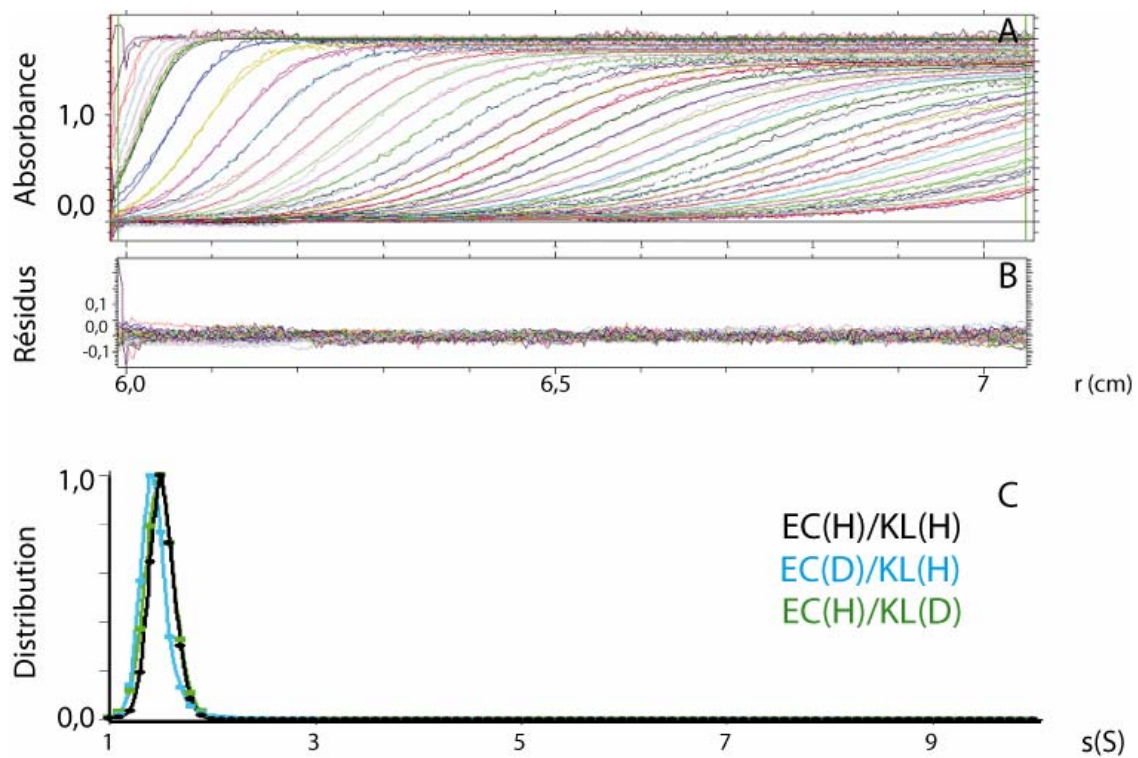


Figure 26 : Ultracentrifugation analytique d'échantillons de complexe EC/KL hydrogéné et partiellement deutérés. A : Superposition de profils de sédimentation expérimentaux d'un échantillon de EC(D)/KL(H) à 13 mg/ml dans 20 mM Tris pH 8, NaCl 150 mM, à 42 000 rpm à 4°C, et des profils simulés et ajustés par le logiciel Sedfit. B : résidus de l'ajustement. C : distribution en fonction de s (S) de EC(D)/KL(H) à 13 mg/ml (en bleu), EC(H)/KL(D) à 10 mg/ml (en vert) et EC(H)/KL(H) à 13 mg/ml (en noir).

IV. Étude par SANS

Les expériences de diffusion aux petits angles de neutrons ont été réalisées sur trois types d'échantillons de complexe EC/KL: des échantillons spécifiquement marqués EC(D)/KL(H) et EC(H)/KL(D), et des échantillons hydrogénés EC(H)/KL(H). Les échantillons ont été purifiés par chromatographie d'affinité sur la résine Ni-NTA® puis à la Streptactin® (cf. chapitre 1/I.), et enfin par chromatographie par exclusion de taille. La pureté des échantillons a été estimée par SDS-PAGE, supérieure à 95% (Figure 25).

IV.1. Caractérisation des échantillons par AUC et DLS

Pour chaque combinaison, 3 échantillons ont été préparés à 3 concentrations différentes : 1,5 , 4 et 10 mg/mL pour EC(H)/KL(H) ; 1,5 , 4 et 13 mg/mL pour EC(D)/KL(H) ; 1,5 , 3 et 10 mg/ml pour EC(H)/KL(D). Ces 9 échantillons ont ensuite été caractérisés par DLS et AUC à 4°C.

Les profils de sédimentation montrent que tous les échantillons sont composés d'une seule espèce apparente (Figure 26).

Cette analyse de sédimentation est cohérente avec la mesure de polydispersité par DLS. En effet, le paramètre de polydispersité de tous les échantillons était inférieur à 30%.

De même que pour les premières analyses du complexe EC/KL hydrogéné, présentées dans ce chapitre au paragraphe II.2, l'ajustement des simili-courbes des profils de sédimentation des 9 échantillons n'est optimal que lorsque le paramètre du rapport des coefficients de friction réel f et de la protéine idéalement sphérique f_0 (f/f_0), est laissé libre d'évoluer, et ce paramètre évolue aussi systématiquement vers des valeurs de 1,6 à 1,8. Le calcul de la masse apparente, avec l'équation 35, à partir du coefficient de sédimentation et du rapport des coefficients de friction donne aussi une masse molaire de 45 kDa environ. Cette masse est plus élevée que la masse d'un complexe EC/KL de stoechiométrie 1:1 (24,8 kDa) et pourrait indiquer la dimérisation du complexe, (EC/KL)₂ de masse 49,6 kDa.

IV.2. Détermination du degré d'oligomérisation du complexe EC/KL

Les courbes de diffusion des 3 échantillons de EC(H)/KL(H), de 1,5 , 4 et 10 mg/mL, ont été mesurées. De ces courbes, la valeur de l'intensité à l'origine $I(0)$ et le rayon de giration R_G ont été extraits. Ces valeurs sont rassemblées dans le tableau 4. La masse de l'objet diffusant a été déterminée à l'aide de l'équation (Jacrot et Zaccai, 1981):

	Conc (mg/ml)	I(0)	M _{app} (kDa)	R _G (Å)
<i>Tampon H₂O</i>		$(990 \pm 5) \cdot 10^{-3}$		
EC(H)/KL(H)	10	$(310 \pm 8) \cdot 10^{-3}$	57 ± 6	55 ± 1
	4	$(91 \pm 7) \cdot 10^{-3}$	42 ± 9	46 ± 3
	1,5	$(41 \pm 7) \cdot 10^{-3}$	50 ± 14	44 ± 7

Tableau 4 : Valeurs de diffusion à l'origine et de rayon de giration d'échantillons de complexe EC/KL hydrogéné. Le calcul de la masse de l'objet diffusant se fait avec la formule citée dans le texte (Eq. 37) et la valeur de diffusion incohérente du tampon H₂O. Les erreurs sur le calcul de la masse prennent en compte une erreur sur l'estimation de la concentration de 10%.

	Conc (mg/ml)	I(0)	I(0)/Conc (10 ⁻³ ml/mg)
EC(D)/KL(H)	13	$(463 \pm 8) \cdot 10^{-3}$	36 ± 4
EC(H)/KL(D)	10	$(352 \pm 7) \cdot 10^{-3}$	35 ± 4

Tableau 5 : Equivalence des intensités à l'origine d'échantillons de complexe EC/KL partiellement deutérés, et complémentaires, dans un tampon 41% D₂O/59% H₂O. L'intensité à l'origine est due exclusivement à la protéine deutérée (en gras). L'erreur sur l'intensité à l'origine normalisée par la concentration prend en compte une erreur de 10% sur l'estimation de la concentration.

Eq. 37 :

$$M_{app} = \frac{1}{f_1} \cdot \frac{I(0)_{exp}}{I(0)_{tampon}} \cdot \frac{1}{c}$$

où M_{app} est la masse moléculaire de l'objet diffusant en solution, c sa concentration, $I(0)_{tampon}$ la diffusion incohérente du tampon et f_1 a été calculée à partir des paramètres de l'expérience (cf. Introduction/B.II.2.1 et l'équation Eq. 45). Ainsi, nous avons déterminé la masse de l'objet diffusant dans les échantillons (Tableau 4) avec $f_1 = 5,5 \cdot 10^{-4} \text{ mol.L.g}^{-2}$.

La masse théorique du complexe EC/KL constitué d'une protéine EC et d'une protéine KL, $(EC/KL)_1$, étant de 24,6 kDa, il est clair que le complexe n'est pas de composition 1/1. La détermination de la masse à partir de la diffusion à l'origine indique que la composition de ce complexe est plutôt 2/2, $(EC/KL)_2$, de masse molaire 49,6 kDa, comme suggéré par les résultats d'ultracentrifugation analytique.

Le rapport entre la quantité de protéine EC et celle de FL a été examiné par la comparaison des intensités à l'origine des courbes des échantillons EC(D)/KL(H) et EC(H)/KL(D). Grâce au mélange isotopique du tampon de ces échantillons, le signal de la protéine hydrogénée du complexe est éteint. Seule est visible la protéine partiellement deutérée. Après normalisation par la concentration (*i.e.* à concentration équivalente), l'intensité à l'origine de diffusion à l'échantillon EC(D)/KL(H) (où seule la protéine EC est visible) est équivalente à l'intensité à l'origine de l'échantillon EC(H)/KL(D) (où seule la protéine KL est visible) (Tableau 5).

Pour une concentration donnée, la contribution de la protéine EC dans le complexe est équivalente à celle de KL. Ces deux protéines sont donc présentes en quantité équimolaire dans le complexe EC/KL.

La masse molaire du complexe EC/KL étant de 50 kDa environ, et les protéines EC et KL en quantité équimolaire, on peut donc conclure que le complexe, dans la gamme de concentration étudiée (de 1 à 10 mg/ml) est constitué de 2 protéines EC et de 2 protéines KL, $(EC/KL)_2$.

IV.3. Détermination de la forme du complexe EC/KL

Le complexe EC/KL est constitué de 2 protéines EC et de 2 protéines KL. Le rayon de giration extrait de la courbe de diffusion est donc celui du complexe $(EC/KL)_2$. Ce rayon est de 50 Å environ (Tableau 4). Ces résultats sont cohérents avec les résultats de DLS qui

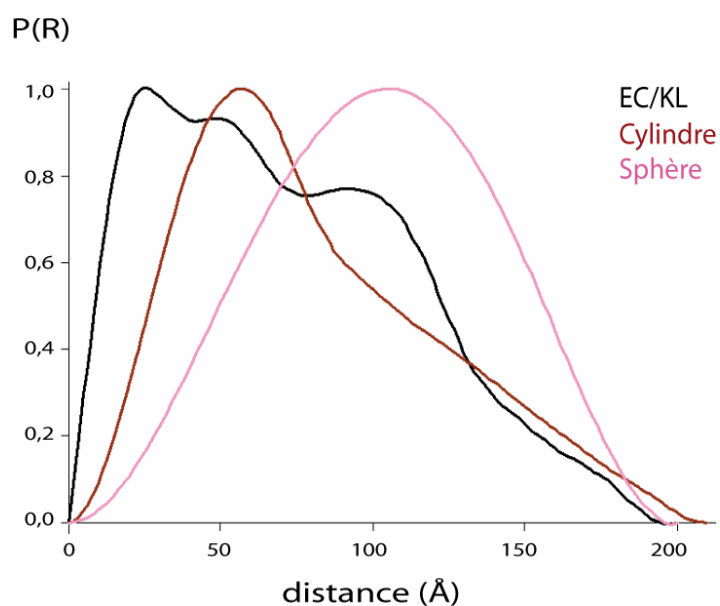


Figure 27 : Comparaison entre la fonction de distributions de distances normalisée du complexe EC/KL et celles d'un cylindre et d'une sphère. La fonction de distribution de paire du complexe EC/KL (en noir), a été estimée à partir de la courbe de diffusion de l'échantillon EC(H)/KL(H) à 10 mg/ml en tampon 100% H₂O. Les fonctions de distribution de la sphère (en rose) et du cylindre (en marron) ont été déterminées à partir d'une sphère de 200 Å et d'un cylindre de 200 Å de long et de 40 Å de diamètre. La distribution normalisée est représentée en fonction de la distance en Å.

indiquaient un rayon de Stokes de 43 Å environ. La valeur du rayon de giration indique que le complexe a des dimensions supérieures à celles qu'aurait une protéine globulaire de 50 kDa. Le complexe est donc non compact. En comparaison, on pourrait citer l'albumine, protéine globulaire (code PDB : 1bm0), de 66 kDa et dont le rayon de giration, calculé avec le logiciel CRYSON (Svergun *et al.*, 1998), vaut 27 Å. La grande taille du complexe peut expliquer pourquoi peu de pics étaient visibles sur le spectre RMN HSQC [^{15}N ; ^1H] d'un échantillon EC(^{14}N)/KL(^{15}N). En effet, un complexe avec un rayon de giration de 50 Å environ est suffisamment grand pour que ses mouvements de diffusion rotationnel dans un solvant soient lents en regard des temps de relaxation de la RMN.

Nous avons tenté d'estimer la fonction de distribution de distances à partir de la courbe de diffusion (Figure 28) avec le logiciel GNOM (Semenyuk et Svergun, 1991). Plusieurs valeurs de distance maximale ($170 \text{ Å} < d_{\text{max}} < 200 \text{ Å}$) ont donné des courbes de distribution de distances comparables. La courbe obtenue pour une distance maximale de 200 Å est présentée en figure 27.

L'estimation de la fonction de distribution de distances montre un profil assez différent du profil d'une molécule sphérique (Figure 27), et plus proche du profil d'une molécule allongée. La forme du complexe semble donc être allongée.

IV.4. Détermination de l'agencement des protéines EC et KL dans le complexe

Nous avons également tenté d'estimer la fonction de distribution de distances des courbes de diffusion des échantillons partiellement deutéré avec le logiciel GNOM (Semenyuk et Svergun, 1991). De même que pour l'échantillon hydrogéné, plusieurs valeurs de distance maximale ($170 \text{ Å} < d_{\text{max}} < 200 \text{ Å}$) ont donné des courbes de distribution de distances, assez similaires. Les courbes obtenues pour une distance maximale de 200 Å sont présentées en figure 29.

Dans les échantillons de complexe partiellement deutéré, seule la protéine deutérée contribue à la courbe de diffusion. Ainsi, la fonction de distribution de distances correspond à la protéine deutérée seule. Ces fonctions de distribution de distances montrent aussi ce profil de protéine allongée, et il est intéressant de remarquer que l'estimation de fonction de distribution de distances des courbes de diffusion des complexe EC(D)/KL(H) et EC(H)/KL(D)

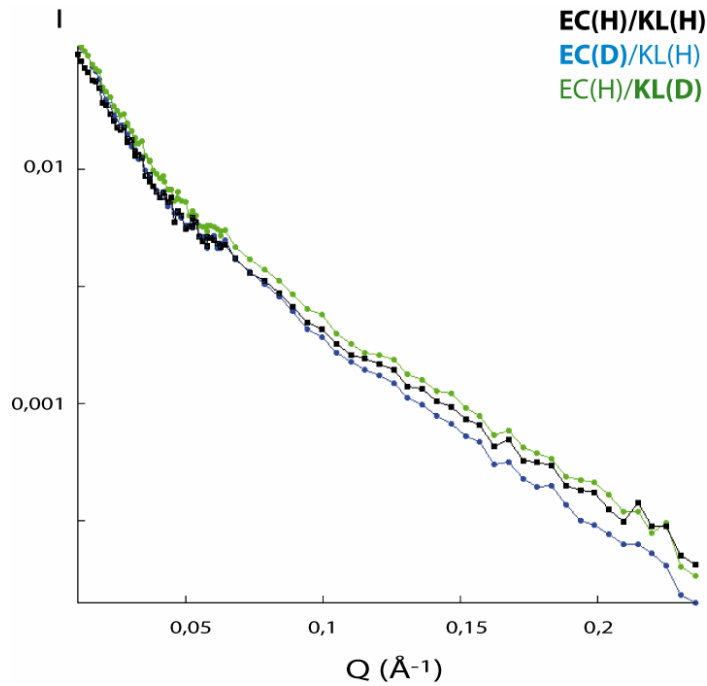


Figure 28 : Comparaison des courbes de diffusion d'échantillons de complexe EC/KL, hydrogéné en tampon 100% H₂O et partiellement deutérés en tampon 41% D₂O/59% H₂O. L'échantillon de complexe hydrogéné EC(H)/KL(H) est en noir, l'échantillon de complexe partiellement deutéré EC(D)/KL(H) est en bleu, l'échantillon de complexe partiellement deutéré EC(H)/KL(D) est en vert, les protéines contribuant à la diffusion sont en gras, dans la légende. L'intensité, en échelle logarithmique, est représentée en fonction Q. Les courbes sont normalisées par la concentration de l'échantillon

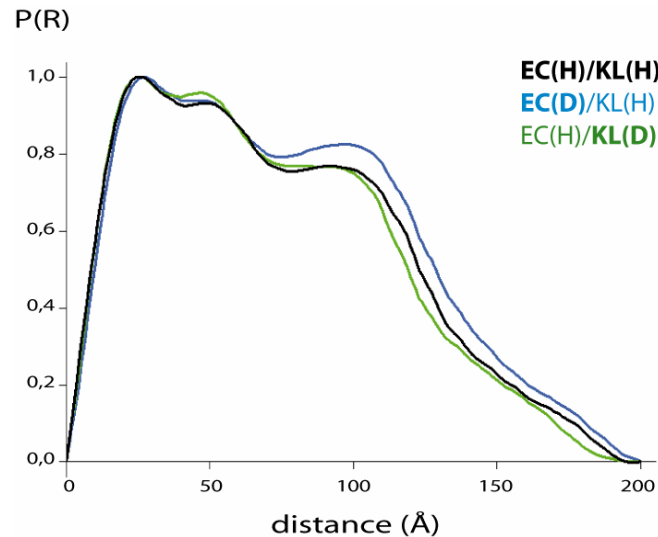


Figure 29 : Comparaison des fonctions de distributions de distances normalisées des complexes EC/KL hydrogéné et partiellement deutérés. La fonction de distribution de paire des protéines EC (en bleu), et celle de KL (en vert), ont été estimées à partir de la courbe de diffusion de l'échantillon EC(D)/KL(H) à 13 mg/ml et de celle de l'échantillon EC(H)/KL(D) à 10 mg/ml, respectivement, en tampon 41% D₂O/ 59 % H₂O. La fonction de distribution de paire du complexe EC/KL (en noir), a été estimée à partir de la courbe de diffusion de l'échantillon EC(H)/KL(H) à 10 mg/ml en tampon 100% H₂O. La distribution normalisée est représentée en fonction de la distance en Å.

dans 41% D₂O/59% H₂O donne un profil très similaire à celle du complexe EC(H)/KL(H) (Figure 29).

Les deux protéines EC et KL, individuellement au sein du complexe, semblent avoir la même forme globale que le complexe lui-même : allongée avec une distribution de distances très similaire à celle du complexe entier. L'allure des courbes de diffusion, très similaires à celle du complexe entier, suggère fortement cette similarité de forme (Figure 28).

De plus, les valeurs de distance maximale sont dans la même gamme que celles pour le complexe entier. La distance maximale de la protéine EC et celle de la protéine KL sont comparables à celle du complexe entier. Ceci indique que les deux protéines de EC et de KL sont réparties tout le long du complexe. Les protéines EC et KL seraient donc parallèles entre elles dans le complexe EC/KL.

V. Étude par SAXS

Pour chercher une concentration où le complexe EC/KL ne serait constitué que d'une protéine EC et d'une protéine KL, (EC/KL)₁, nous avons décidé de poursuivre l'étude par diffusion aux petits angles de rayons X. En effet, le flux important de rayons X de la ligne ID02 de l'ESRF permet de mesurer la diffusion d'échantillons moins concentrés par rapport à la diffusion aux petits angles de neutrons. Nous avons donc mesuré la courbe de diffusion d'échantillons de complexe EC/KL, dilué à partir d'un échantillon à 1 mg/mL. De ces courbes, la valeur de l'intensité à l'origine $I(0)$ a été extraite. Ces valeurs sont rassemblées dans le tableau 6. La masse de l'objet diffusant a été déterminée à l'aide de l'équation :

Eq. 38 :

$$M_{app} = \frac{1}{f_2} \frac{I(0)_{exp}}{c}$$

où M est la masse moléculaire de l'objet diffusant en solution et c sa concentration et $f_2 = (83 \pm 1) \cdot 10^{-6} \text{ mol.L.g}^{-2}$, déterminée à partir de la diffusion d'un échantillon de lysozyme (cf. Introduction/B.II.2.1 et l'équation Eq. 46).

Moins l'échantillon de EC/KL est concentré, plus la masse apparente diminue. Pour un échantillon de EC/KL de 0,2 mg/mL, la masse calculée est de 26 kDa, ce qui correspond à la masse d'un complexe formé d'une protéine EC et d'une protéine KL. A ces concentrations, le

	Conc (mg/ml)	I(0)	M _{app} (kDa)
EC/KL	1 (40 μ M)	$(4,48 \pm 0,10) \cdot 10^{-3}$	54 ± 7
	0,8 (32 μ M)	$(2,94 \pm 0,07) \cdot 10^{-3}$	45 ± 6
	0,6 (24 μ M)	$(2,00 \pm 0,06) \cdot 10^{-3}$	40 ± 6
	0,4 (16 μ M)	$(1,16 \pm 0,05) \cdot 10^{-3}$	35 ± 5
	0,2 (8 μ M)	$(0,43 \pm 0,05) \cdot 10^{-3}$	26 ± 6

Tableau 6 : Valeurs d'intensité de diffusion à l'origine d'échantillons de complexe EC/KL. Le calcul de la masse de l'objet diffusant se fait avec la formule citée dans le texte (Eq. 38), dont le paramètre de calibration a été extrait de la valeur de diffusion à l'origine du lysozyme. Les erreurs sur le calcul de la masse prennent en compte une erreur sur l'estimation de la concentration de 10%.

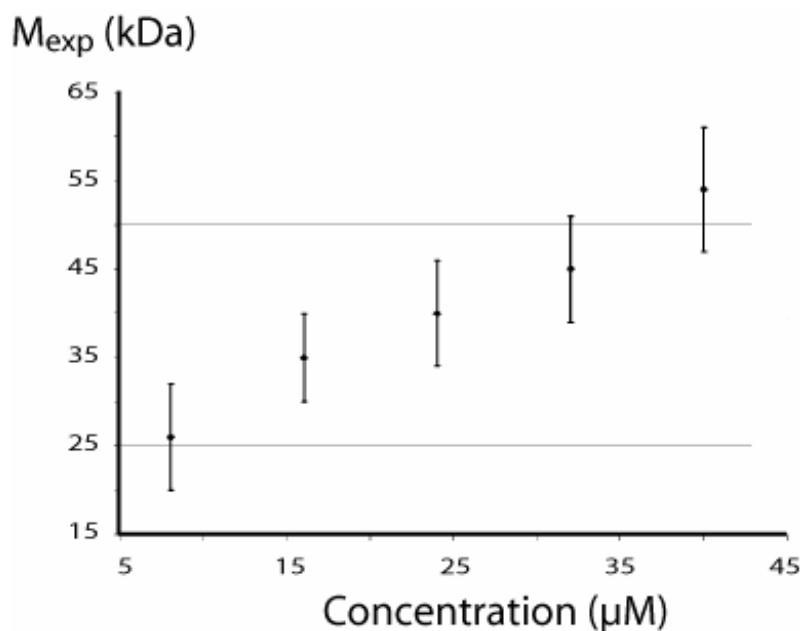


Figure 30 : Evolution de la masse apparente du complexe EC/KL en fonction de la concentration. La masse apparente des échantillons de EC/KL déterminée à partir de l'intensité de diffusion à l'origine est représentée en fonction de la concentration en micromolaire des échantillons, de 8 μ M (0.2 mg/ml) à 40 μ M (1 mg/ml).

complexe E/K est encore formé puisque sa constante de dissociation est de l'ordre du nanomolaire (Chao et al., 1996), et qu'un échantillon de complexe EC/KL à 0,2 mg/mL correspond à une concentration de 8 μ M. L'échantillon à 0,2 mg/ml est donc constitué de complexe (EC/KL)₁. Le complexe EC/KL est donc en équilibre entre la forme (EC/KL)₁ et la forme (EC/KL)₂,

$$2(EC/KL)_1 \leftrightarrow (EC/KL)_2 \quad K_D = \frac{[(EC/KL)_1]^2}{[(EC/KL)_2]}$$

La constante de dissociation est la concentration à laquelle la concentration de (EC/KL)₂ est égale à celle de (EC/KL)₁.

L'échantillon concentré à 8 μ M semble composé de (EC/KL)₁ majoritairement puisque la masse apparente de cet échantillon est de 26 kDa. L'échantillon concentré à 40 μ M semble composé de (EC/KL)₂ majoritairement puisque la masse apparente de cet échantillon est de 54 kDa. La concentration de l'échantillon où la concentration de (EC/KL)₁ est égale à celle de (EC/KL)₂, est donc comprise entre 8 et 40 μ M. Cette constante de dissociation est donc de l'ordre de 25 μ M. En effet, la masse apparente de l'échantillon à 24 μ M est de 40 kDa, ce qui correspond à une moyenne entre la masse de (EC/KL)₁ et celle de (EC/KL)₂ (Figure 30). Une estimation plus fine de cette constante n'a pas pu être faite sur ces données, du fait des grandes erreurs sur les mesures.

Nous avons donc trouvé une condition où le complexe EC/KL n'est constitué que d'une protéine EC et d'une protéine KL : à très basse concentration. Malheureusement cette condition ne peut être exploitée pour réaliser une modélisation basse résolution du complexe. En effet, aux valeurs de Q au-delà de la zone de Guinier, la diffusion d'un échantillon à cette concentration est à la limite de la détection de l'instrument et la courbe obtenue est trop bruitée pour permettre une modélisation. La structure à basse résolution du complexe (EC/KL)₁ n'a pas pu être modélisée.

Chapitre 2. La partie extracellulaire de DivIB de *S. pneumoniae*

Nous avons mené une étude structurale sur le domaine extracellulaire de la protéine DivIB de *S. pneumoniae*. Le but de cette étude était de déterminer la structure tridimensionnelle de cette protéine, à la meilleure résolution possible.

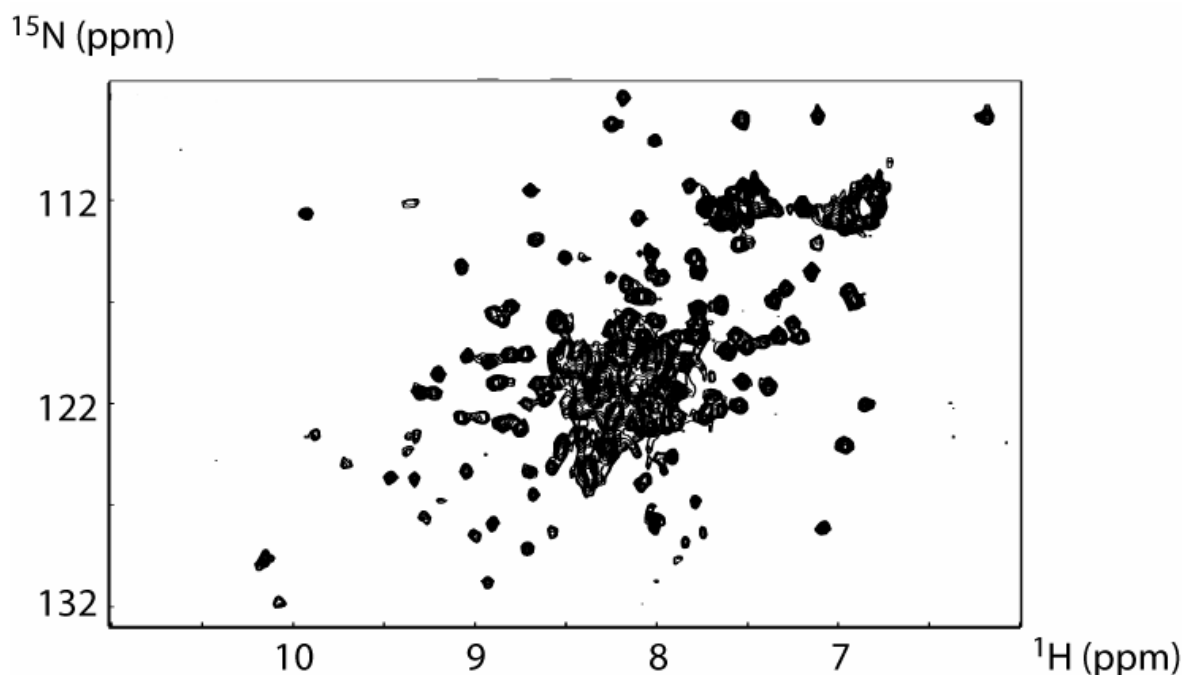


Figure 31 : Spectre HSQC [^{15}N ; ^1H] de la partie extracellulaire de DivIB de *S. pneumoniae*. L'échantillon était à une concentration de 140 μM dans une solution tamponnée : 20 mM Tris pH 8, 150 mM NaCl, 90% H_2O /10% D_2O .

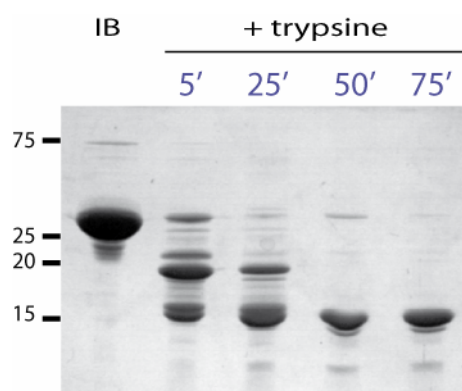


Figure 32 : Protéolyse limitée de la partie extracellulaire de DivIB de *S. pneumoniae*. SDS-PAGE coloré au bleu de Coomassie d'échantillons des digestions tryptiques de IB, incubées à 37°C, pendant 5, 25, 50 et 75 minutes. Le témoin (IB) est un échantillon sans trypsine, en première piste.

I. Analyse préliminaire du domaine extracellulaire de DivIB (RMN et protéolyse limitée)

I.1. Analyse de la structuration du domaine extracellulaire de DivIB

Pour une analyse préliminaire du domaine extracellulaire de DivIB, de 247 acides aminés (de la sérine 150 à l'arginine 396), appelé IB (cf. Introduction/C.), nous avons commencé l'étude par RMN. Le spectre des fréquences de résonance des protons d'un échantillon de IB montre quelques pics à des valeurs de déplacements chimiques inférieurs à 0 ppm. Cette zone du spectre correspondant aux fréquences de résonance des protons des groupes méthyles, ces pics révèlent la présence d'un cœur hydrophobe qui indique que la protéine est structurée, au moins en partie. Nous avons donc poursuivi par l'enregistrement d'un spectre HSQC [^{15}N ; ^1H]. Pour enregistrer ce spectre, un échantillon de IB, marqué ^{15}N a été préparé.

Le spectre montre des pics dispersés, et une zone de pics non résolus dans la région 120 ppm ; 8 ppm (Figure 31). Les protons amides d'acides aminés non structurés résonnent dans cette zone de fréquence. On peut donc en déduire qu'une partie de la protéine n'est pas structurée. D'autre part, on compte une centaine de pics isolés. Comme dans ce type de spectre, un pic correspond à un proton amide, on peut en déduire qu'une centaine d'acides aminés sont structurés dans cet échantillon. La protéine IB était constituée de 247 acides aminés, elle est donc en partie structurée, avec environ 100 acides aminés structurés, et environ 150 non structurés.

Afin d'isoler biochimiquement la région structurée, nous avons poursuivi par des expériences de protéolyse limitée par la trypsine. Les mélanges réactionnels (trypsine:IB = 1:100) ont été incubés à 37°C pendant des temps variables. Les mélanges réactionnels ont ensuite été analysés par SDS-PAGE et coloration en bleu de Coomassie (Figure 32).

La protéine IB, de 28 kDa, est protéolysée partiellement (Figure 32). Rapidement (après 5 min d'incubation), deux bandes d'intensité majoritaire apparaissent, vers 20 et 15 kDa. Au bout de 50 min, la bande à 20 kDa disparaît. La bande de 15 kDa résiste à la digestion tryptique après 75 min d'incubation. Ainsi, dans le domaine extracellulaire de DivIB, on trouve un domaine structuré, résistant à la trypsine de 15 kDa environ. Le reste de la protéine IB, invisible sur SDS-PAGE, est digéré en fragments trop petits pour être détectés.

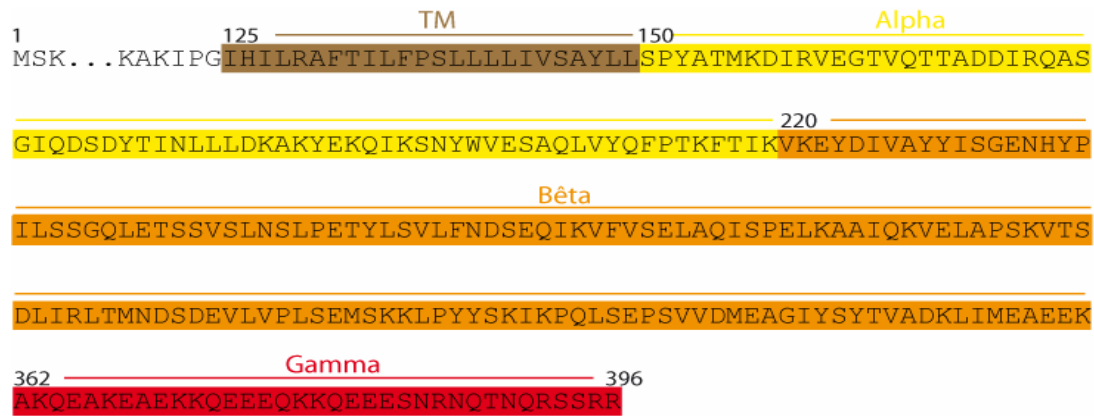


Figure 33 : Séquence primaire de DivIB de *S. pneumoniae*. En marron le segment transmembranaire (TM) ; en jaune le domaine α ; en orange, le domaine β ; en rouge, le domaine γ . Les positions au-dessus des acides aminés N-ter des domaines α , β , et γ sont indiqués, ainsi que la méthionine 1 et l'arginine C-ter 396. Le domaine cytoplasmique n'est pas représenté en entier, il est symbolisé par des pointillés.

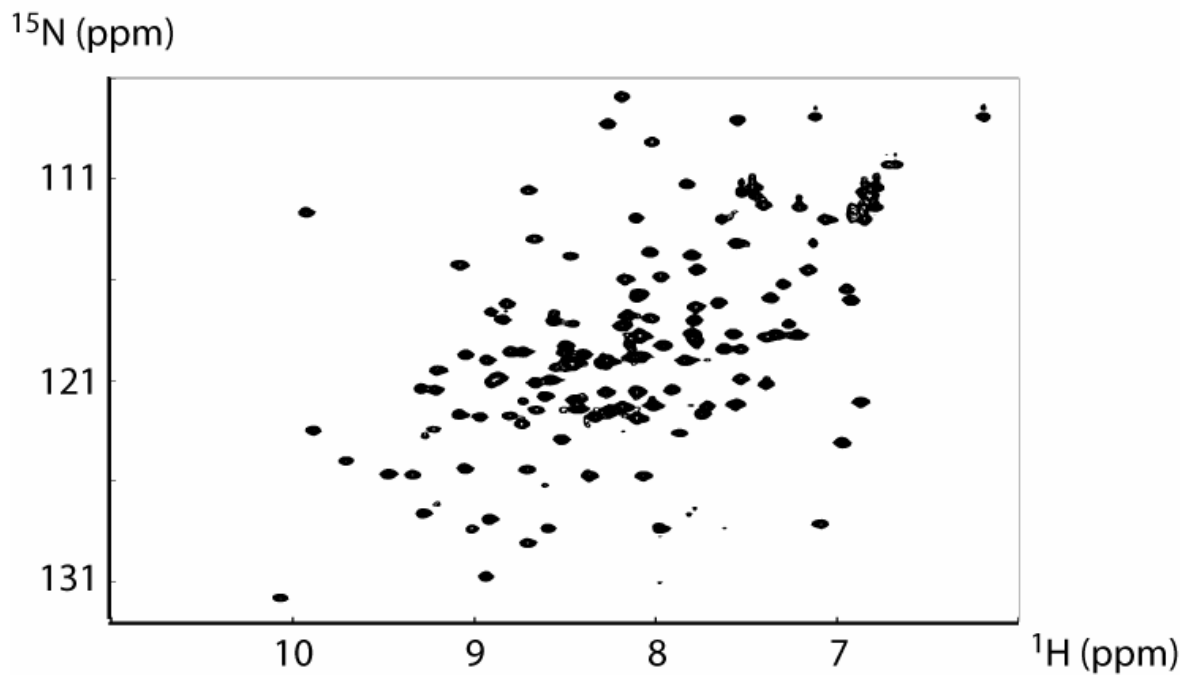


Figure 34 : Spectre HSQC [^{15}N ; ^1H] du domaine β de DivIB de *S. pneumoniae*. L'échantillon est à une concentration de 270 μM dans une solution tamponnée : 20 mM Tris pH 8, 150 mM NaCl, 90% H_2O /10% D_2O .

Ces résultats sont cohérents avec les spectres de RMN qui indiquent que la partie extracellulaire de DivIB est partiellement structurée, avec environ 100 acides aminés structurés qui pourraient correspondre au domaine résistant à la protéolyse de 15 kDa environ.

I.2. Délimitation de trois domaines de la partie extracellulaire de DivIB

Une analyse en spectrométrie de masse et un séquençage des acides aminés N-terminaux sur un échantillon de ce domaine résistant à la trypsine (préparé par protéolyse limitée d'un échantillon de protéine IB), ont permis de définir les bornes de ce domaine. La masse de ce domaine mesuré par spectrométrie de masse était de $15\,842 \pm 1$ Da et les premiers acides aminés de sa séquence étaient VKEYDIVA. Ces résultats ont permis de définir un domaine, de la valine 220 à la lysine 361 (141 acides aminés), qui est appelé le domaine β (Figure 33). D'autre part, l'analyse par spectrométrie de masse n'a montré la présence que d'une seule espèce, confirmant que le reste de la protéine IB, i.e. les régions en N-ter et en C-ter de ce domaine β , a été protéolysée. Le domaine extracellulaire de DivIB peut donc être considéré comme constitué de trois domaines : α , β , et γ . Le domaine α , de la sérine 150 à la lysine 219, compte 70 résidus ; β , le domaine résistant à la trypsine, 141 résidus et γ , de l'alanine 362 à l'arginine 396, 35 résidus. Ces limites sont définies par la protéolyse, elles dépendent donc de la structuration de la protéine, mais aussi de la spécificité de la protéase utilisée, la trypsine dans ce cas, qui hydrolyse les liaisons peptidiques en C-ter des lysines et arginines.

La partie extracellulaire de DivIB jusqu'à présent appelée IB, est donc constituée de trois domaines α , β , et γ , et pour plus de cohérence avec la dénomination en domaines, cette protéine sera par la suite appelée $\alpha\beta\gamma$.

Un échantillon de β a été préparé par digestion tryptique à partir d'un échantillon de $\alpha\beta\gamma$ (ou IB) marqué ^{15}N . Le spectre HSQC [^{15}N ; ^1H] de cet échantillon de β a été enregistré et montre environ 120 pics dispersés (Figure 34).

Le domaine β contient 140 acides aminés, dont 8 prolines qui ne possèdent pas de proton amide, et qui n'apparaissent donc pas dans ce spectre. L'analyse du spectre indique que 120 acides aminés sont bien structurés, au moins, car les prolines peuvent aussi être dans des parties structurées bien qu'elles ne soient pas visibles sur ce spectre.

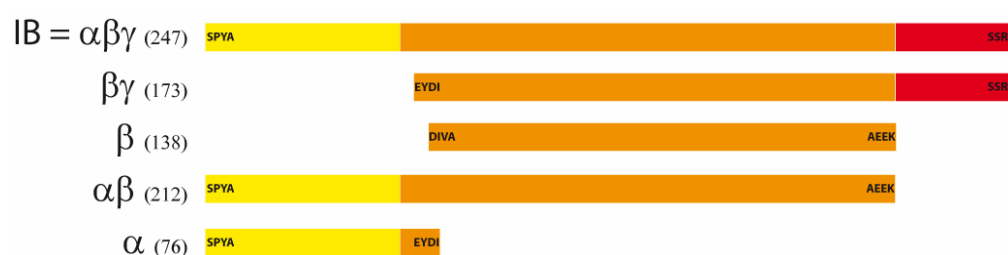


Figure 35 : Schéma des protéines variantes de DivIB de *S. pneumoniae*. En jaune le domaine α ; en orange, le domaine β ; en rouge, le domaine γ . En face des constructions, est indiqué le nom utilisé, et entre parenthèses, le nombre de résidus des constructions. Les premiers et les derniers acides aminés des variants sont indiqués en début et fin des constructions.

La superposition du spectre du domaine extracellulaire entier, $\alpha\beta\gamma$ (ou IB), et de celui de β montre que les pics dispersés du spectre de $\alpha\beta\gamma$ (ou IB) se retrouvent dans le spectre du domaine β seul. Seuls une dizaine de pics du spectre de $\alpha\beta\gamma$ (ou IB) ne se retrouvent pas dans le spectre de β . On peut donc attribuer la contribution structurée du spectre de $\alpha\beta\gamma$ (ou IB) à β , et la contribution non structurée à α et γ . La superposition indique aussi que le domaine β est structuré indépendamment de α et γ , puisque les déplacements chimiques de ses protons amides ne sont pas influencés par la présence ou l'absence des domaines α et γ .

La partie extracellulaire recombinante de DivIB de *S. pneumoniae* de 28 kDa, est donc une protéine partiellement structurée, constituée de trois domaines. La partie structurée est constituée d'un domaine central, β de 15 kDa, indépendant des autres domaines. Le premier domaine, proche de la membrane, α , ne semble pas structuré, ainsi que le domaine comprenant les résidus C-terminaux, γ .

II. Étude structurale des domaines de la partie extracellulaire de DivIB

II.1. Production des variants tronqués de la partie extracellulaire de DivIB

Nous avons désiré ne pas restreindre l'étude au seul domaine structuré β . C'est pourquoi nous avons élaboré plusieurs constructions génétiques, qui correspondent à des découpages de la protéine $\alpha\beta\gamma$ (ou IB) : le domaine α seul, les domaines $\alpha\beta$, le domaine β seul, et les domaines $\beta\gamma$ (Figure 35). Nous avons donc voulu poursuivre l'étude avec ces 4 protéines, en plus de la protéine $\alpha\beta\gamma$ (ou IB).

Les séquences codants pour chaque construction ont été fusionnés en 3' du gène codant pour la glutathion S-transférase (GST) pour permettre la purification des protéines par chromatographie d'affinité sur glutathion. Un site de reconnaissance de la protéase Tobacco Etched Virus (TEV) a été inséré, afin de pouvoir cliver la chaîne peptidique entre la GST et la protéine d'intérêt. Des tests d'expression, puis de purification ont été menés sur ces cinq constructions.

	α	β	$\alpha\beta$	$\beta\gamma$	$\alpha\beta\gamma$ (ou IB)
Rendement de purification par litre de culture	3 mg	20 mg	6 mg	16 mg	8 mg

Tableau 7 : Rendement de purification des différents variants de la partie extracellulaire de DivIB de *S. pneumoniae*. Les protéines ont été purifiées par chromatographie d'affinité sur glutathion Sepharose. Le domaine GST n'a pas été clivé.

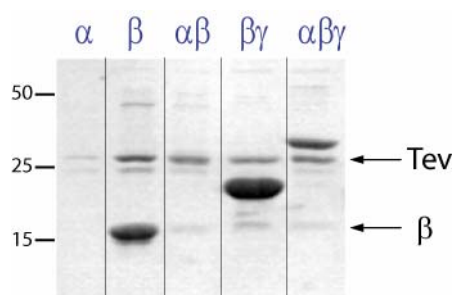


Figure 36 : Bilan des purifications des différentes constructions de la partie extracellulaire de DivIB de *S. pneumoniae*. SDS-PAGE 15% coloré au bleu de Coomassie. Les échantillons ont été purifiés sur colonne de chromatographie d'affinité, et déposés après un clivage par la protéase TEV de l'étiquette GST.

	β	$\beta\gamma$	$\alpha\beta\gamma$ (ou IB)
Masse théorique	15 kDa	20 kDa	28 kDa
Masse apparente	23 ± 2 kDa	32 ± 3 kDa	53 ± 4 kDa

Tableau 8 : Masses apparentes des protéines β , $\beta\gamma$ et $\alpha\beta\gamma$ (ou IB). Ces masses ont été déterminées par chromatographie d'exclusion de taille calibrée.

Les protéines ont été exprimées dans *E. coli*, en production classique en Erlen-Meyer, et nous nous sommes concentrés sur la fraction soluble des protéines. Les résultats de purification des protéines non clivées (avec la fusion GST, M = 28kDa) sont reportés dans le tableau 7.

La faible quantité de protéine α , $\alpha\beta$ et $\alpha\beta\gamma$ (ou IB) purifiée est en partie due à la faible proportion de protéine soluble. La fraction soluble de ces protéines est d'autre part sensible à la protéolyse : des bandes de dégradation sont visibles sur les premiers gels d'analyse (SDS-PAGE). Le domaine α n'est rapidement plus détectable sur SDS-PAGE coloré au bleu de Coomassie, et $\alpha\beta$ est rapidement dégradé en β (quelques jours à 4°C). Il n'a pas été possible d'en purifier suffisamment pour permettre une identification par spectrométrie de masse, ou une quantification. Sans précaution particulière (manipulation à 4°C, ajout d'inhibiteurs de protéases), la construction $\alpha\beta\gamma$ (ou IB) se dégrade également en $\beta\gamma$ en quelques jours, ce qui est cohérent avec l'instabilité apparente de α .

Les variants β et $\beta\gamma$, ont des niveaux d'expression élevé, sont majoritairement solubles, et sont stables face à la dégradation.

La figure 36 présente le bilan des purifications de ces 5 protéines après le clivage à la protéase TEV du domaine GST. Sur le gel présenté, on voit clairement l'absence de la protéine α , et la dégradation de la protéine $\alpha\beta$ en β , 5 jours après la purification.

Les résultats des tests d'expression et de purification ont permis de retenir les variants β et $\beta\gamma$, en plus de la construction initiale $\alpha\beta\gamma$ (ou IB). Les variants α et $\alpha\beta$ ont été abandonnés.

II.2. Caractérisation de l'auto-association des protéines

Pour tenter de déterminer leur degré d'oligomérisation, les trois protéines $\alpha\beta\gamma$ (ou IB), $\beta\gamma$, et β , ont été analysées par chromatographie d'exclusion de taille. Le profil d'élution des trois protéines montre un pic d'élution, à partir duquel une masse apparente a été déterminée, grâce à une calibration de la colonne de chromatographie (Tableau 8).

Les protéines β et $\beta\gamma$ semblent monomériques. La masse déterminée est un peu supérieure à la masse théorique. La détermination de la masse par cette technique repose sur un modèle de protéine globulaire. En effet, une protéine allongée sera éluée comme une protéine globulaire de masse légèrement plus grande. Or, le domaine β d'*E. coli* est de forme

allongée (van den Ent *et al.*, 2008), il est donc possible et attendu que les protéines β et $\beta\gamma$ soient aussi de forme allongée, ce qui expliquerait ces masses surestimées.

La masse déterminée de la protéine $\alpha\beta\gamma$ (ou IB) est aussi supérieure à la masse théorique. Nous avons donc voulu vérifier la présence de multimères par une autre technique, l'ultracentrifugation analytique. L'analyse des profils de vitesse de sédimentation d'un échantillon de $\alpha\beta\gamma$ (ou IB) par le logiciel Sedfit (Schuck, 2000) montre une distribution en fonction des coefficients de sédimentation composée majoritairement (63 %) d'une espèce ayant un coefficient de sédimentation de 1,74 S. La masse de cette espèce estimée par le logiciel était de 27,5 kDa environ (avec un R_s estimé à 35,1 Å et Equation 34). La masse du monomère étant de 28 kDa, il semble que l'échantillon soit composé majoritairement de monomères de $\alpha\beta\gamma$ (ou IB).

Cependant, l'analyse montre aussi des formes de masse moléculaire plus grande. D'ailleurs dans la chromatographie par exclusion de taille, une fraction de l'échantillon se retrouvait dans le volume exclu. Il est probable que $\alpha\beta\gamma$ (ou IB) forme des oligomères, que nous avons essayé d'éliminer afin d'isoler la fraction monomérique pour les expériences suivantes.

II.3. Étude structurale du domaine β par RMN

La structure des domaines β d'*E. coli* et de *G. stearothermophilus* étant connue (Robson et King, 2006; van den Ent *et al.*, 2008), la résolution de la structure de β de *S. pneumoniae* n'a pas été jugée prioritaire. Cependant, l'attribution des fréquences de résonance des atomes de la chaîne principale de la protéine a été réalisée sur un échantillon de la protéine β , marqué ^{13}C et ^{15}N . Cette attribution a été exploitée dans le cadre d'expériences sur le complexe ternaire que β forme avec le complexe EC/KL, rapportée au chapitre 3.

Les domaines α et γ n'étant pas structurés a priori, une étude structurale classique par RMN sur ces domaines n'a pas pu être réalisée. L'étude a donc été poursuivie par diffusion aux petits angles de rayons X.

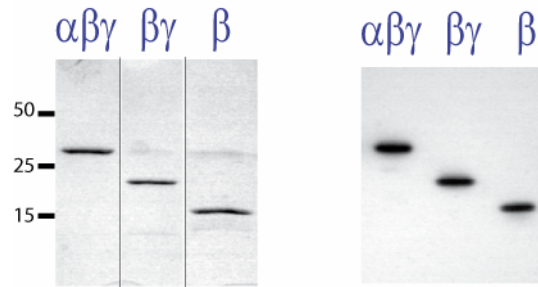


Figure 37 : Contrôle de la qualité des échantillons utilisés pour les expériences de SAXS. SDS-PAGE coloré au bleu de Coomassie, à gauche; SDS-PAGE immuno-révéle à droite, avec des anticorps polyclonaux anti-DivIB de *S. pneumoniae*.

	Conc (mg/ml)	I(0)	M _{app} (kDa)	R _G (Å)
β	10,5	$(13,8 \pm 0,1) \cdot 10^{-3}$	16 ± 2	$20,4 \pm 0,3$
	8,5	$(10,4 \pm 0,1) \cdot 10^{-3}$	15 ± 2	$20,0 \pm 0,4$
	4,5	$(5,9 \pm 0,1) \cdot 10^{-3}$	16 ± 2	$21,3 \pm 0,5$
	1,8	$(2,5 \pm 0,1) \cdot 10^{-3}$	17 ± 2	$21,7 \pm 1,0$

Tableau 9 : Valeurs de diffusion à l'origine et de rayon de giration d'échantillons du domaine β de *S. pneumoniae*. La masse de l'objet diffusant a été déterminée avec la formule citée dans le texte (eq. 38), dont le paramètre de calibration, f_2 , a été extrait de la valeur de diffusion à l'origine du lysozyme. Les erreurs sur le calcul de la masse prennent en compte une erreur sur l'estimation de la concentration de 10%.

II.4. Etude structurale par SAXS

II.4.1. Préparation des échantillons

Pour obtenir une grande quantité de protéines, un bioréacteur a été utilisé. Le protocole de culture des bactéries et d'expression des protéines en bioréacteur a été élaboré au laboratoire, par Benoit Gallet (LIM/IBS) et moi-même. Ce protocole permet d'augmenter jusqu'à 5 fois la masse des bactéries récoltées par rapport à une production en Erlen-Meyer classique.

La pureté des échantillons étant essentielle pour cette étude, nous avons mis au point un protocole de purification en trois étapes. Une première étape de purification a été réalisée par chromatographie d'affinité (sur glutathion Sepharose), puis une seconde par chromatographie échangeuse d'ions, et la dernière, par chromatographie d'exclusion de taille pour isoler la fraction monomérique des protéines. La pureté des protéines a été analysée par SDS-PAGE et estimée supérieure à 95%. L'identité des protéines est vérifiée par western blot et spectrométrie de masse (Figure 37). L'homogénéité des protéines a été vérifiée par PAGE en conditions non dénaturantes.

II.4.2. Etude du domaine β

Une gamme de concentration, de 10,5 mg/mL à 1,8 mg/mL, a été préparée. Les courbes de diffusion de quatre échantillons de β , à des concentrations différentes, ont été mesurées. De ces courbes, la valeur de l'intensité à l'origine $I(0)$ et le rayon de giration R_G ont été extraits. Ces valeurs sont rassemblées dans le tableau 9. La masse de l'objet diffusant a été déterminée à partir de l'intensité à l'origine et de la concentration de l'échantillon, en utilisant la relation suivante :

Eq. 38 :

$$M_{app} = \frac{1}{f_2} \cdot \frac{I(0)_{exp}}{c}$$

où M est la masse moléculaire de l'objet diffusant en solution et c sa concentration (cf. Introduction/B.II.2.1). Le facteur f_2 a été déterminé par une calibration à l'aide d'un échantillon de lysozyme de concentration connue (cf. *Eq. 46*). Les valeurs des masses déterminées, en considérant les échantillons monodisperses, sont données dans le tableau 9.

	Nbre de résidus	M _{théo} (kDa)	R _G (Å)
β_{Sp}	138	15,3	20,8
β_{Ec}	134	15,1	17,4
β_{Gs}	115	12,9	15,3

Tableau 10 : Comparaison des domaines β de *S. pneumoniae*, *E. coli* et *G. stearothermophilus* : taille et rayon de giration. Les rayons de giration des domaines β d'*E. coli* (Ec) et de *G. stearothermophilus* (Gs) ont été extraits des simili-courbes de diffusion calculée par CRY SOL. Celui du domaine β de *S. pneumoniae* (Sp) a été extrait de la courbe expérimentale.

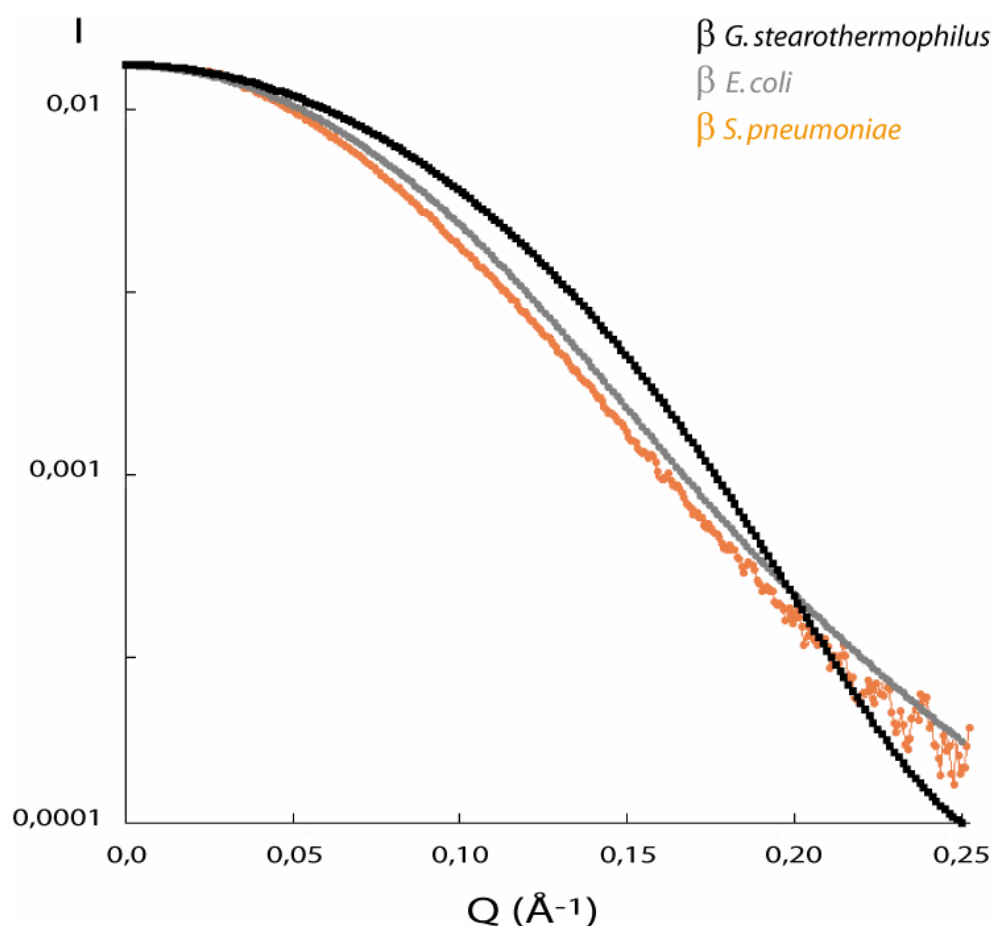


Figure 38 : Comparaison des courbes de diffusion aux petits angles du domaine β de différentes espèces de bactéries. La courbe de diffusion de β de *S. pneumoniae* (en orange) est expérimentale. Les courbes de diffusion de β de *E. coli* (en gris) et *G. stearothermophilus* (en noir) sont calculées à partir des structures déposées dans la PDB, 2vh1 et 1yr1 respectivement, avec le logiciel CRY SOL. L'intensité, en échelle logarithmique, est représentée en fonction de Q.

La masse théorique de β étant de 15,3 kDa, on peut en effet considérer que les échantillons contiennent que des monomères de β . Le rayon de giration extrait de la courbe peut donc être attribué à la protéine. Ce rayon est d'environ $20,8 \pm 1,0 \text{ \AA}$. Pour comparaison, les rayons de giration calculés du domaine β de *E. coli* et de celui de *G. stearothermophilus* sont indiqués dans le tableau 10. Ces rayons de giration ont été extraits des simili-courbes de diffusion de ces protéines, calculées par le logiciel CRY SOL (Svergun, 1995) à partir de leur structure, (code PDB : 2vh1 et 1yr1 respectivement).

Le rayon de giration du domaine β de *S. pneumoniae* (β_{Sp}) est plus grand que celui des autres espèces. La forme du domaine β de *S. pneumoniae* semble donc moins globulaire. Il est toutefois important de noter que le domaine β de *G. stearothermophilus* (β_{Gs}) est plus petit de 24 résidus.

En comparant la courbe expérimentale de β de *S. pneumoniae* aux simili-courbes calculées par CRY SOL de β_{Gs} et β_{Ec} , on remarque qu'elle est plus proche de la courbe de β de *E. coli*. On remarque aussi la différence entre les courbes de β de *E. coli* et de *G. stearothermophilus* (Figure 38).

Comme attendue, la courbe de diffusion du domaine β de *G. stearothermophilus* a un rayon de giration plus petit et décroît plus lentement que celle du domaine β d'*E. coli*. Ceci indique bien que le domaine β de *G. stearothermophilus* est plus petit, en taille, que celui d'*E. coli*, comme on peut le voir en comparant les structures tridimensionnelles (cf. Introduction / A.V.4.3.2 ; Figure 12).

La courbe expérimentale de diffusion du domaine β de *S. pneumoniae* indique un rayon de giration un peu plus grand que celui d'*E. coli*, mais leur décroissance est très similaire, et presque superposable à partir de $0,1 \text{ \AA}^{-1}$. La comparaison des courbes indique donc que la forme du domaine β de *S. pneumoniae* est similaire à celle du domaine β d'*E. coli*, et plus proche de celle du domaine β d'*E. coli* que de celle du domaine β de *G. stearothermophilus*.

	Conc (mg/ml)	I(0)	M _{app} (kDa)	R _G (Å)
$\beta\gamma$	7,5	$(15,3 \pm 0,1) \cdot 10^{-3}$	25 ± 2	$26,6 \pm 0,7$
	3,7	$(7,7 \pm 0,1) \cdot 10^{-3}$	25 ± 2	$27,3 \pm 1,6$
	1,5	$(3,3 \pm 0,1) \cdot 10^{-3}$	27 ± 3	$25,6 \pm 3,1$
	0,8	$(1,8 \pm 0,1) \cdot 10^{-3}$	28 ± 3	$29,2 \pm 4,5$

Tableau 11 : Valeurs de l'intensité à l'origine et du rayon de giration d'échantillons de la protéine $\beta\gamma$ de *S. pneumoniae*. Le calcul de la masse de l'objet diffusant se fait avec la formule citée dans le texte (Eq.38), dont le paramètre de calibration a été extrait de la valeur de diffusion à l'origine du lysozyme. Les erreurs sur le calcul de la masse prennent en compte une erreur sur l'estimation de la concentration de 10%.

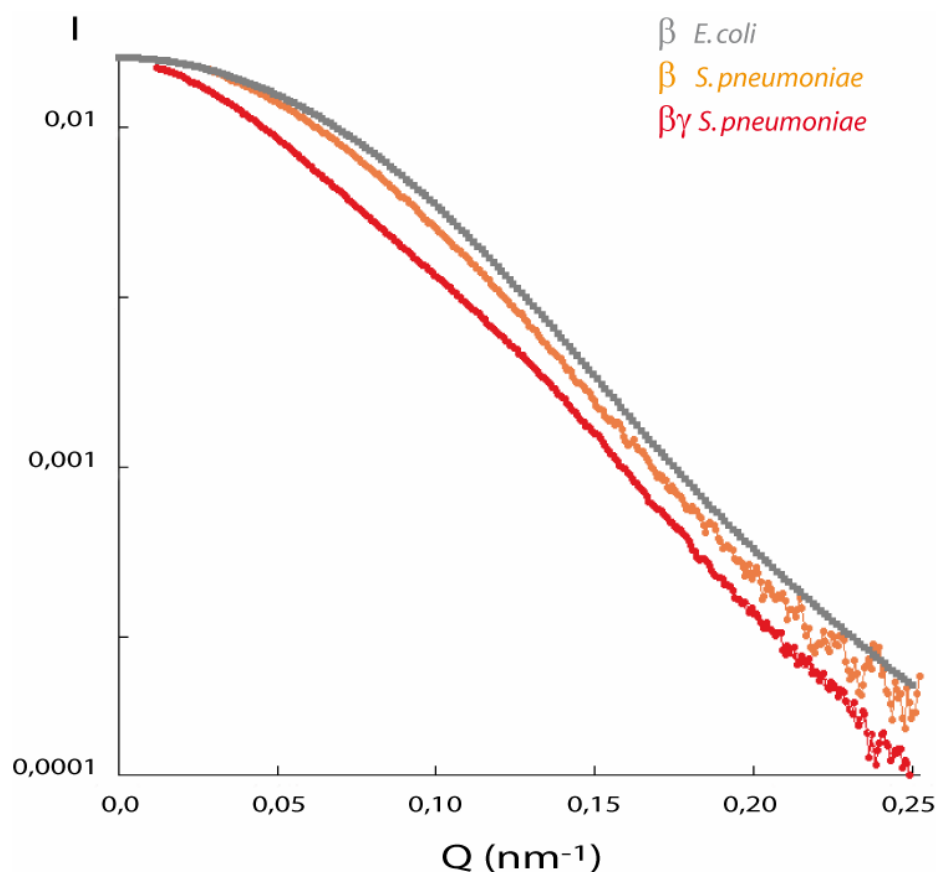


Figure 39 : Comparaison des courbes de diffusion des domaines $\beta\gamma$, et β de *S. pneumoniae*, avec la simili-courbe de diffusion du domaine β d'*E. coli*. Les courbes de diffusion de $\beta\gamma$ (en rouge) et de β (en orange) de *S. pneumoniae* sont expérimentales. La courbe de diffusion de β d'*E. coli* (en gris) est calculée à partir de la structure 2vh1 déposée dans la PDB, avec le logiciel CRY SOL. L'intensité, en échelle logarithmique, est représentée en fonction de Q . Les intensités sont normalisées par la masse et par la concentration de l'échantillon.

II.4.3. Etude de la protéine $\beta\gamma$

De même que pour le domaine β , une gamme de concentration a été préparée pour $\beta\gamma$. Les courbes de diffusion de quatre échantillons de $\beta\gamma$, à des concentrations différentes, ont été mesurées. De ces courbes, ont été extraits la valeur de l'intensité à l'origine $I(0)$ et le rayon de giration R_G . Ces valeurs sont rassemblées dans le tableau 11.

Avec la même formule et le même facteur f_2 , que pour β , nous avons déterminé la masse apparente de l'objet diffusant dans les échantillons de $\beta\gamma$, considérant que les échantillons étaient monodisperses.

La masse théorique de $\beta\gamma$ est 20 kDa, on remarque que la masse déterminée à partir de l'intensité diffusée à l'origine est supérieure de 25 % à la masse théorique. Cependant, elle est suffisamment éloignée de la masse d'un dimère potentiel, pour que l'on puisse considérer l'échantillon comme étant composé de monomères de $\beta\gamma$. Le rayon de giration extrait de la courbe peut donc être attribué à la protéine $\beta\gamma$. Ce rayon est de 27 ± 2 Å.

Une comparaison de la courbe de diffusion de $\beta\gamma$ avec la courbe et simili-courbe de diffusion du domaine β de *S. pneumoniae*, et de celui d'*E. coli*, montre une différence aux petits angles mais une grande similarité aux valeurs de Q supérieures à $0,13 \text{ Å}^{-1}$ (Figure 39).

La différence entre les courbes aux petits angles vient du rayon de giration de $\beta\gamma$, plus grand que celui de β , de 7 angströms environ. La décroissance plus rapide de la courbe de $\beta\gamma$ indique aussi que $\beta\gamma$ est de plus grande taille que β , comme attendu. En effet, $\beta\gamma$ (173 résidus) contient 35 résidus de plus que β (138 résidus).

La fonction de distribution de distances de la courbe de diffusion de $\beta\gamma$ a été calculée avec le logiciel GNOM (Semenyuk et Svergun, 1991). La meilleure estimation de cette fonction a été trouvée pour une distance maximale de la protéine, d_{\max} , de 90 Å. Une modélisation *ab initio* de la forme de $\beta\gamma$ a été réalisée avec le logiciel DAMMIN (Svergun, 1999), à partir de cette fonction de distribution de distances. La modélisation a été réalisée à partir d'une sphère de 90 Å de diamètre, remplie de 9688 simili-atomes de 1,9 Å de rayon. La modélisation a été réalisée indépendamment 5 fois, et les 5 modèles sont très similaires (Figure 40).

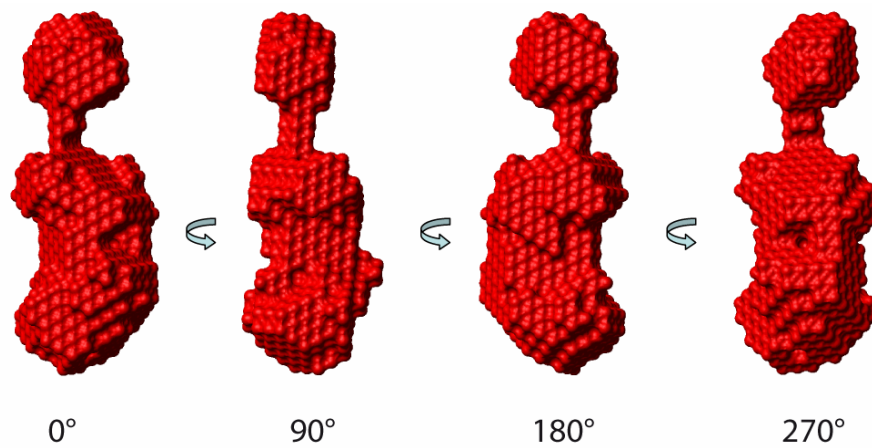


Figure 40 : Modèle *ab initio* à basse résolution des domaines $\beta\gamma$ de DivIB de *S. pneumoniae*. Les différentes représentations du modèle correspondent à des rotations de 90° autour de l'axe vertical

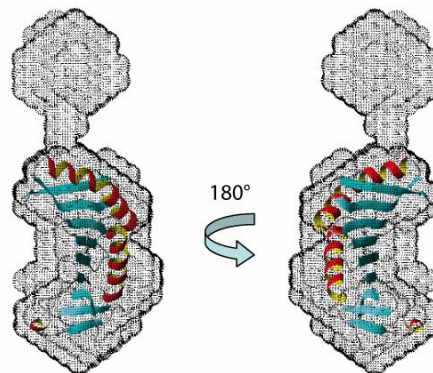


Figure 41 : Comparaison de l'enveloppe du modèle des domaines $\beta\gamma$ de *S. pneumoniae* (obtenue par modélisation *ab initio*), avec la structure cristallographique du domaine β d'*E. coli*. La superposition a été effectuée à la main.

Le modèle basse résolution du domaine $\beta\gamma$ montre deux domaines : un grand domaine d'environ 45 Å de long, légèrement courbé, et un domaine globulaire d'environ 25 Å de diamètre. Le domaine globulaire est accroché à une extrémité du domaine courbé.

Il semble que le modèle basse résolution du domaine légèrement courbé de $\beta\gamma$ de *S. pneumoniae* enveloppe bien la structure du domaine β d'*E. coli* (Figure 41). Ce domaine courbé pourrait donc correspondre au domaine β de *S. pneumoniae*. D'autre part, le volume du petit domaine globulaire est supérieur au volume d'un peptide globulaire de 20 résidus, mais inférieur au volume d'un peptide globulaire de 50 résidus. Ce petit domaine pourrait donc correspondre aux 35 résidus du domaine γ . Les résidus du domaine γ ne sont pas protégés de la protéolyse de la trypsine et ne semblent donc pas structurés. Du point de vue du modèle *ab initio*, le petit domaine pourrait correspondre à une moyenne des conformations échantillonnées par le domaine γ (non structuré et prenant donc plusieurs conformations).

Ainsi, la forme du domaine β de *S. pneumoniae* serait très proche de celle de β d'*E. coli*, et le domaine γ , à l'extrémité C-terminale de β , non structuré.

	Conc (mg/ml)	I(0)	M _{app} (kDa)	R _G (Å)
$\alpha\beta\gamma$ (ou IB)	8,5	$(79,7 \pm 0,4) \cdot 10^{-3}$	113 ± 3	$54,9 \pm 0,5$
	6,8	$(58,3 \pm 0,5) \cdot 10^{-3}$	103 ± 3	$54,4 \pm 0,9$
	3,5	$(21,6 \pm 0,3) \cdot 10^{-3}$	76 ± 3	$51,6 \pm 0,9$
	1,4	$(7,7 \pm 0,2) \cdot 10^{-3}$	68 ± 4	$49,3 \pm 1,7$
	0,7	$(3,3 \pm 0,2) \cdot 10^{-3}$	59 ± 4	$49,2 \pm 3,3$

Tableau 12 : Valeurs de la diffusion à l'origine et du rayon de giration d'échantillons de la protéine $\alpha\beta\gamma$ (ou IB) de *S. pneumoniae*. La masse de l'objet diffusant a été déterminée avec l'équation 38, dont le paramètre de calibration f_2 , a été extrait de la valeur de diffusion à l'origine du lysozyme. Les erreurs sur le calcul de la masse prennent en compte une erreur sur l'estimation de la concentration de 10%.

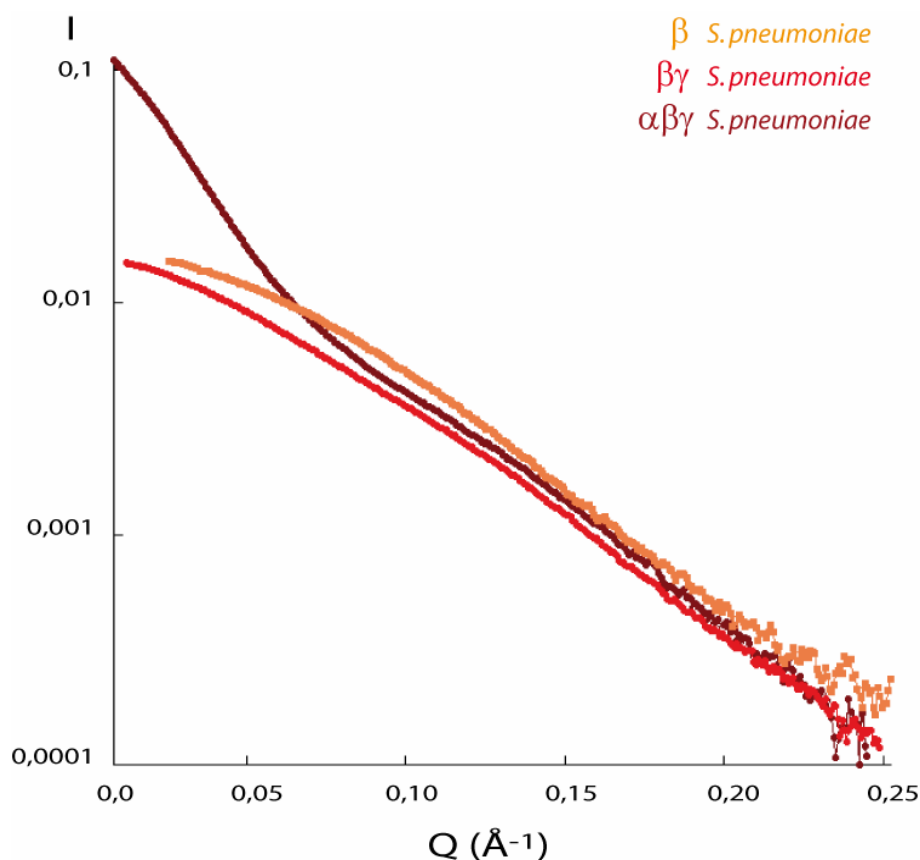


Figure 42 : Comparaison des courbes de diffusion des différents variants de la partie extracellulaire de DivIB de *S. pneumoniae*. La courbe expérimentale de l'échantillon $\alpha\beta\gamma$ (ou IB) à 8,5 mg/ml (en marron), est comparée à celle de $\beta\gamma$ (en rouge), et de β (en orange). L'intensité, en échelle logarithmique, est représentée en fonction de Q . Les intensités sont normalisées par la masse et la concentration des échantillons.

II.4.4. Etude de la protéine $\alpha\beta\gamma$ (ou IB)

Une gamme de concentrations de $\alpha\beta\gamma$ (ou IB) a été préparée. Les courbes de diffusion de cinq échantillons de $\alpha\beta\gamma$ (ou IB), à des concentrations différentes, ont été mesurées. De ces courbes, ont été extraites la valeur de l'intensité à l'origine $I(0)$ et le rayon de giration R_G . Ces valeurs sont rassemblées dans le tableau 12.

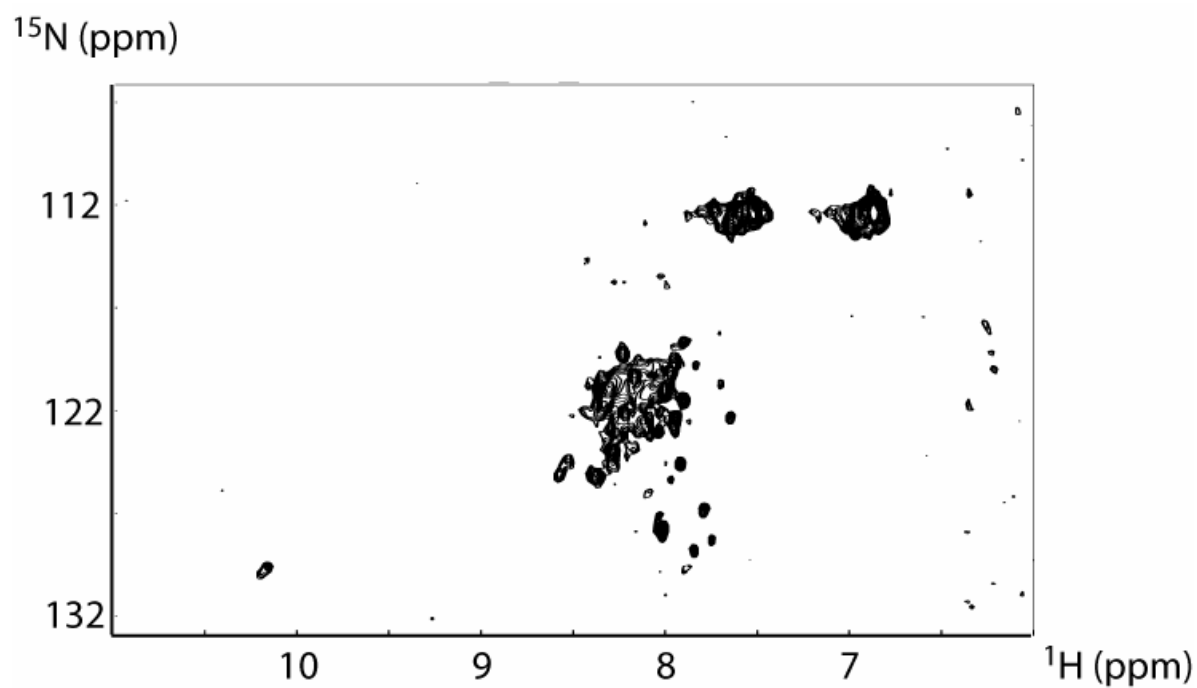
La masse déterminée est la masse moyenne des objets diffusant en solution. La masse théorique du monomère de $\alpha\beta\gamma$ (ou IB) est 28 kDa. Les divers échantillons ne contiennent donc pas uniquement des monomères de $\alpha\beta\gamma$ (ou IB). En effet, malgré les précautions prises pour tenter d'isoler la fraction monomérique, l'intensité à l'origine met en évidence la présence de multimères ou d'agrégats de $\alpha\beta\gamma$ (ou IB). Le rayon de giration extrait des courbes n'est donc pas attribuable à une protéine monomérique. La comparaison des courbes de diffusion de β , $\beta\gamma$ et $\alpha\beta\gamma$ (ou IB) montre bien cette présence de grosses particules dans l'échantillon de $\alpha\beta\gamma$ (ou IB) (Figure 42).

Aux grands angles, au dessus de $0,15 \text{ \AA}^{-1}$, les courbes sont assez similaires. L'effet de la présence des multimères ou des agrégats de $\alpha\beta\gamma$ (ou IB) est très net aux petits angles, où la courbe de $\alpha\beta\gamma$ (ou IB) diverge, pour atteindre une intensité de diffusion à l'origine dix fois plus grande que celle des deux autres courbes.

Ainsi, l'étude par diffusion aux petits angles des échantillons de $\alpha\beta\gamma$ (ou IB) n'a pas pu donner d'autres informations structurales.

Chapitre 3. L'interaction entre le complexe contraint des domaines extracellulaires de DivIC et FtsL, EC/KL, et le domaine extracellulaire de DivIB de *S. pneumoniae*

Nous avons mené une étude de l'interaction entre le complexe contraint des domaines extracellulaires de DivIC et de FtsL, EC/KL, et le domaine extracellulaire de DivIB. L'objectif était de mieux connaître l'interface entre ces protéines. D'autre part, nous avons également mené une étude structurale sur le complexe des trois protéines, afin d'obtenir la structure de ce complexe trimérique et l'arrangement des trois protéines dans le complexe.



*Figure 43 : Spectre HSQC [^{15}N ; ^1H] de la partie extracellulaire de DivIB de *S. pneumoniae* complexée avec le complexe EC/KL. L'échantillon était à une concentration de 110 μM dans une solution tamponnée : 20 mM Tris pH 8, 150 mM NaCl, 90% H_2O /10% D_2O .*

I. Etude par RMN sur la partie extracellulaire de DivIB complexée avec EC/KL

I.1. Analyse préliminaire

Nous avons commencé l'étude par RMN. Le spectre HSQC d'un échantillon du domaine extracellulaire de DivIB, $\alpha\beta\gamma$ (ou IB) marqué ^{15}N (cf. Chap. 2/I.1.), auquel a été ajouté le complexe contraint EC/KL non marqué, a été enregistré pour observer des éventuelles variations de déplacements chimiques des protons de la protéine $\alpha\beta\gamma$ (ou IB). Le rapport stoechiométrique final entre $\alpha\beta\gamma$ et EC/KL d'environ de 1:2 et leur concentration, permettaient d'espérer qu'une majorité des molécules de $\alpha\beta\gamma$ (ou IB) soit complexée avec EC/KL.

Le spectre de $\alpha\beta\gamma$ (ou IB) complexée montre quelques pics résolus et une zone de pics non résolus dans la région (120 ppm ; 8 ppm) (Figure 43). On ne retrouve pas tous les pics des fréquences de résonance de la protéine $\alpha\beta\gamma$ (ou IB) seule (Figure 31 ; chapitre 2 /I.1). Cependant, on ne remarque pas une augmentation franche du nombre de pics dans la zone non résolue, par rapport au spectre de la protéine $\alpha\beta\gamma$ (ou IB) seule. La protéine $\alpha\beta\gamma$ (ou IB) ne s'est donc pas déstructurée au contact de EC/KL.

Les pics ont disparu dans le bruit de fond. Cette disparition est induite par un changement du comportement hydrodynamique de la protéine, et notamment de sa diffusion rotationnelle. Ceci indiquerait que la protéine $\alpha\beta\gamma$ (ou IB) est bien dans un complexe dont le temps de diffusion rotationnelle est plus grand. La protéine $\alpha\beta\gamma$ (ou IB) serait donc bien complexée avec EC/KL.

Il semble également que l'interaction se fasse via le domaine β . En effet, les pics des acides aminés du domaine β ont disparu, et les pics des domaines α et γ , non structurés, ne semblent pas avoir été influencés par l'interaction avec EC/KL. Les parties non structurés, flexibles, ne semblent donc pas participer à l'interaction et conservent leur propriété angulaire, due à leur flexibilité.

Pour vérifier cette hypothèse, le spectre HSQC du domaine β auquel le complexe EC/KL non marqué a été ajouté, a été enregistré. Dans ce spectre, on ne distingue plus que quelques pics très peu distincts du bruit de fond.

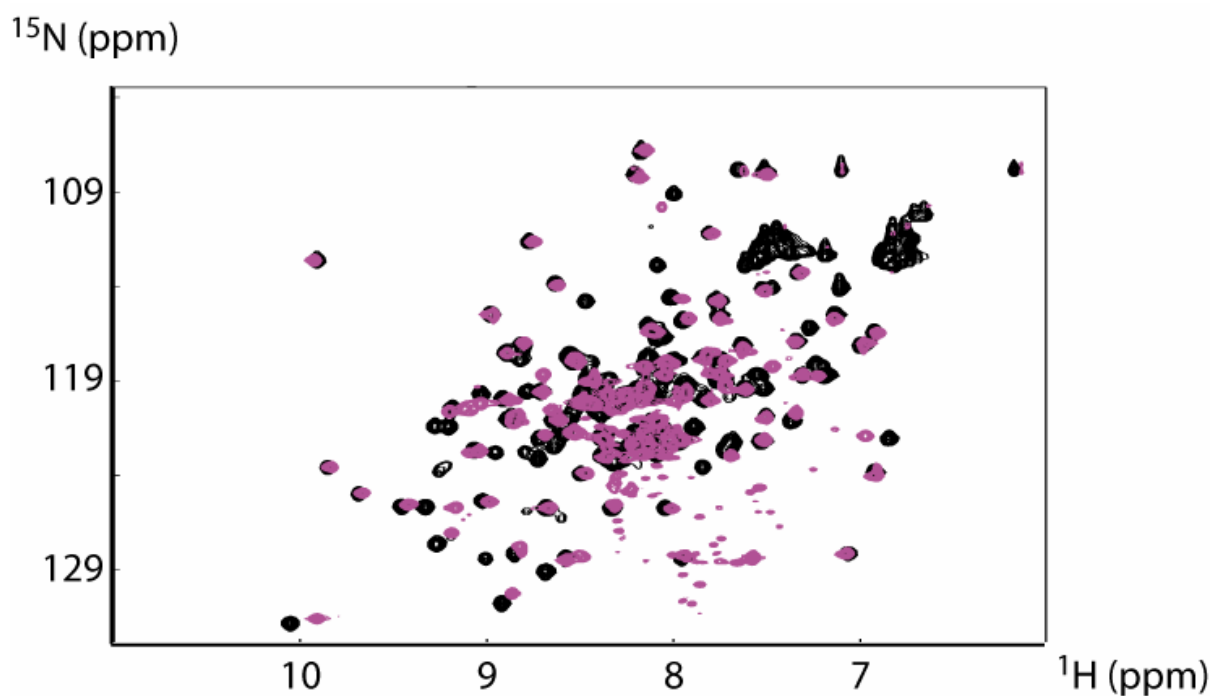


Figure 44 : Superposition des spectres HSQC [^{15}N ; ^1H] du domaine β (^{15}N ; ^2H) de DivIB de *S. pneumoniae*, seul et complexé avec EC/KL. Le spectre de β seul est en noir, complexé avec EC/KL, en violet. L'échantillon de la protéine β était à une concentration de 400 μM , dans une solution tamponnée : 20 mM Tris pH 8, 150 mM NaCl, 90% H_2O /10% D_2O .

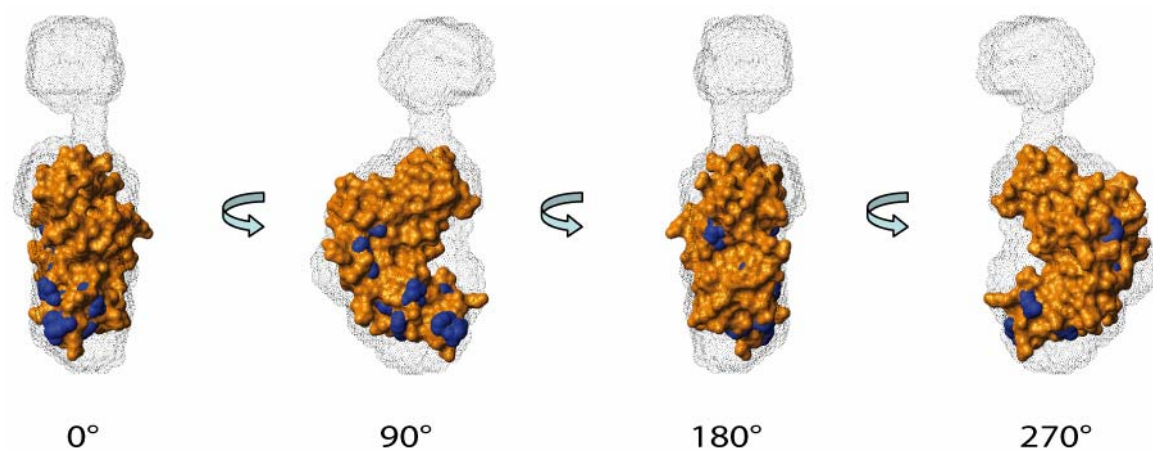


Figure 45 : Modèle de la zone d'interaction entre le domaine β et le complexe EC/KL sur la structure du domaine β d'*E. coli*. Les atomes des acides aminés du domaine β d'*E. coli* correspondants aux acides aminés du domaine β de *S. pneumoniae* impliqués dans l'interaction avec EC/KL sont colorés en bleu. Tous les autres atomes sont colorés en orange. Une représentation transparente du modèle de structure du domaine $\beta\gamma$ de *S. pneumoniae* obtenu par SAXS (en noir) enveloppe la structure du domaine β d'*E. coli* pour visualiser la zone d'interaction sur le modèle basse résolution.

La diminution des intensités des pics de la protéine β (par rapport au spectre de la protéine seule) montre que le complexe se forme en l'absence des domaines α et γ . Cette observation confirme l'interaction via le domaine β .

Les pics résiduels, superposables aux pics du spectre de la protéine β seule, correspondent donc à la fraction libre de la protéine β , non complexée à EC/KL.

La disparition des pics dans le bruit de fond étant due à la relaxation transversale, l'étude a été poursuivie sur un échantillon deutéré du domaine β , pour minimiser cette relaxation et tenter d'observer le spectre du complexe.

I.2. Détermination de l'épitope du domaine β impliqué dans l'interaction

Le spectre HSQC [^{15}N ; ^1H] d'un échantillon du domaine β (^{15}N , ^2H) (cf. Chap. 2/II.3), auquel a été ajouté un échantillon de complexe contraint EC/KL non marqué, a été enregistré, pour observer des éventuelles variations de déplacements chimiques des protons du domaine β . Il était espéré que le marquage partiel au deutérium permettrait de diminuer suffisamment les effets de la relaxation pour que les pics de la protéine β complexée avec EC/KL se distinguent du bruit de fond. Cela fut le cas et des variations de déplacement chimique ont pu être observées (Figure 44).

Les variations de déplacements chimiques indiquent que les protons correspondants sont influencés par l'interface avec EC/KL. L'attribution des atomes de la protéine β (cf. Chap. 2/II.3) a permis d'identifier les acides aminés dont les protons amides sont impliqués dans l'interaction avec EC/KL. Ces acides aminés sont la tyrosine 229, la sérine 231, la tyrosine 236, le glutamate 245, la valine 249, la sérine 250, la leucine 254, l'alanine 293, la leucine 301, la valine 312, la leucine 313 et la sérine 320. Il est important de remarquer que cette liste n'est sans doute pas exhaustive. En effet, n'ont été repérés que les acides aminés dont les variations de déplacements chimiques étaient nettement résolues. De plus, seuls les acides aminés dont le proton amide est influencé par l'interaction ont été distingués.

En utilisant l'alignement des séquences primaires des protéines DivIB de *S. pneumoniae* et d'*E. coli* présenté en figure 59 (cf. discussion/II.), ainsi que la structure du domaine β d'*E. coli* (code PDB : 2vh1), une représentation de la zone d'interface a pu être modélisée (Figure 45).

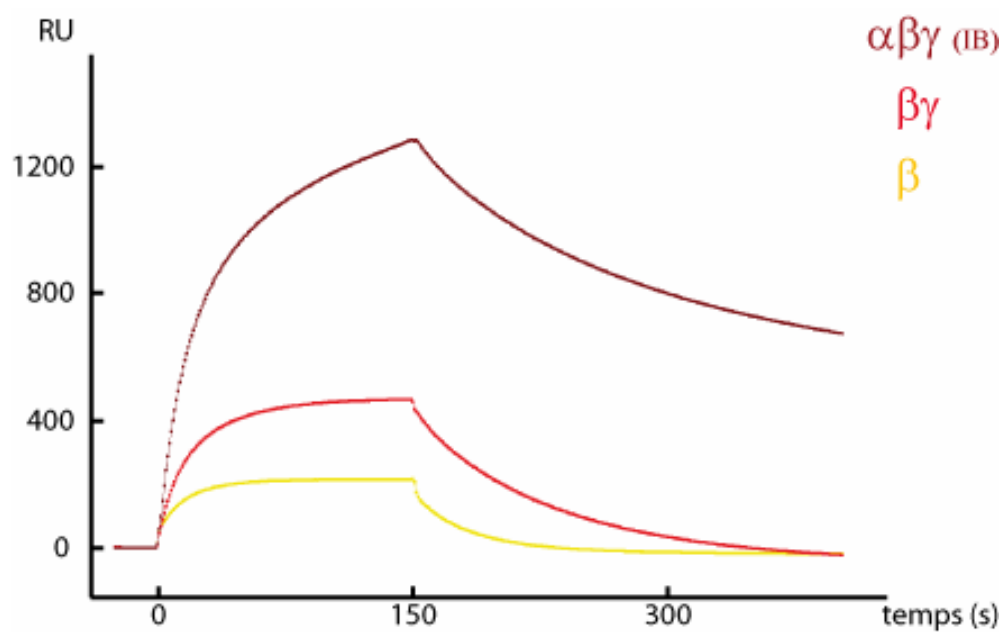


Figure 46 : Comparaison de l'interaction avec le complexe EC/KL des protéines $\alpha\beta\gamma$ (ou IB), $\beta\gamma$ et β . Les courbes d'interaction des protéines β à 1,25 μ M (en jaune), $\beta\gamma$ à 1,15 μ M (en rouge) et $\alpha\beta\gamma$ à 1 μ M (en marron), avec (His6)EC/(His6)KL fixé sur un biosenseur, ont été suivies par résonance plasmonique de surface, avec un instrument Biacore®.

II. Cartographie des domaines d'interaction du domaine extracellulaire de DivIB et du complexe EC/KL par analyse d'interaction de biomolécules par SPR

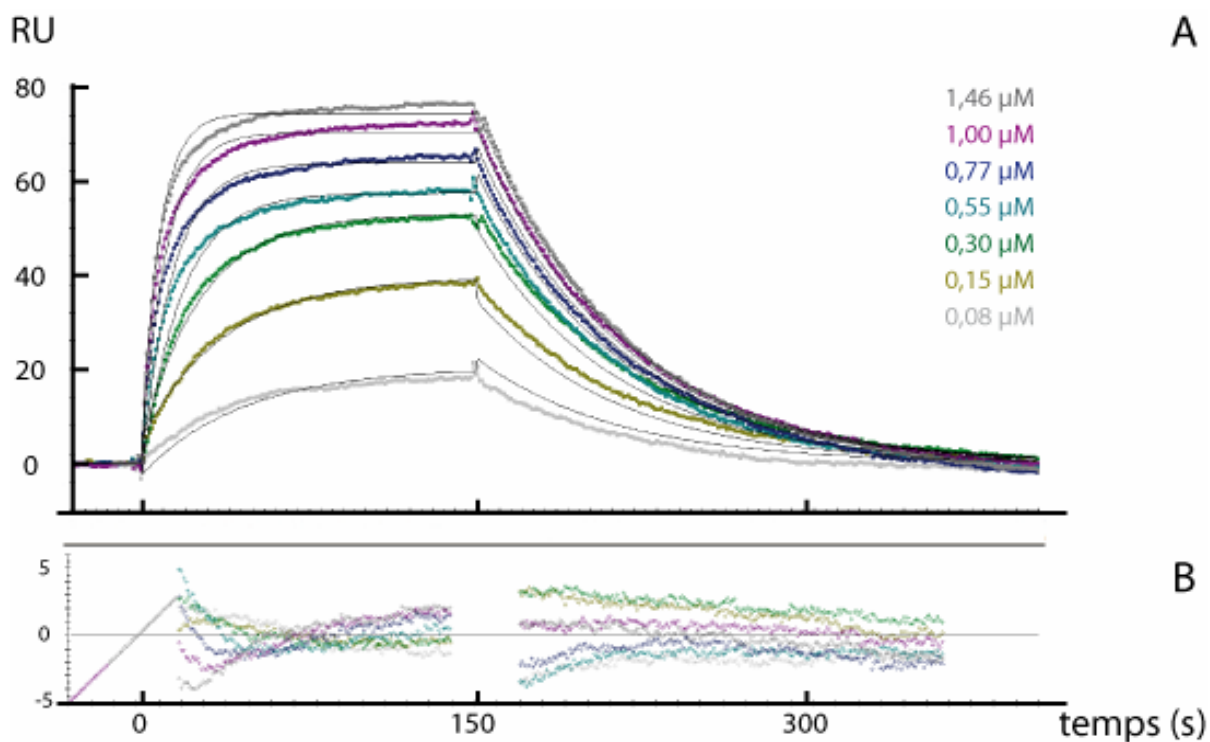
Des expériences d'analyse d'interaction de biomolécules par résonance plasmonique de surface ont été réalisées entre la protéine $\alpha\beta\gamma$ (ou IB), et les variants $\beta\gamma$ et β , et le complexe contraint EC/KL, avec un instrument Biacore.

II.1. Mesure de l'interaction entre les domaines de la partie extracellulaire de DivIB et le complexe EC/KL

Le but de cette étude était de mesurer l'association entre la partie extracellulaire de DivIB, $\alpha\beta\gamma$ (ou IB), avec le complexe contraint EC/KL. Pour cela, nous avons, dans un premier temps, comparé l'interaction des différents variants de la partie extracellulaire de DivIB : β , $\beta\gamma$ et $\alpha\beta\gamma$ (ou IB, la partie extracellulaire entière).

Le complexe contraint EC/KL est fixé via les étiquettes poly-histidine des protéines, sur le bio senseur dont la surface est saturée avec des ions de nickel (dans ces expériences ce sont les constructions (His6)EC et (His6)KL qui ont été utilisées). Les analytes, i.e. les protéines β , $\beta\gamma$ et $\alpha\beta\gamma$, sont additionnés au tampon qui est injecté en continu dans les pistes du système micro-fluidique, et l'association avec EC/KL est mesurée. La dissociation est aussi mesurée en stoppant l'addition des protéines analytes dans le tampon.

Dans le cas de la protéine $\alpha\beta\gamma$, le signal d'association-dissociation n'est pas simple à interpréter, en comparaison à celui observé avec les protéines β , $\beta\gamma$ (Figure 46). Pour ces 2 dernières protéines, on observe nettement l'association jusqu'à une saturation, puis une dissociation totale du complexe avec EC/KL. La courbe d'interaction de la protéine $\alpha\beta\gamma$ est peut-être plus difficile à interpréter du fait d'une probable agrégation de cette protéine. En effet, nous avons vu au chapitre 2 que cette protéine formait des oligomères, et que les échantillons n'étaient pas monodisperses, rendant impossible une étude structurale par diffusion aux petits angles. Il semble que cette protéine ne soit pas non plus adaptée pour une étude par résonance plasmonique de surface. Nous avons donc décidé de poursuivre les mesures d'affinité pour EC/KL avec la protéine $\beta\gamma$.



k_A (1/M.s)	k_D (1/s)	K_A (1/M)	K_D (M)	χ^2
$7,34 \cdot 10^4$	0,0161	$4,56 \cdot 10^6$	$0,219 \cdot 10^{-6}$	2,26

Tableau 13 : Valeurs des constantes de vitesses d'association et de dissociation, des constantes d'association et de dissociation du complexe formé par la protéine $\beta\gamma$ et le complexe EC/KL, et de l'erreur de l'estimation. Les courbes expérimentales ont été ajustées avec un modèle d'association de Langmuir.

Les courbes d'interaction de la protéine $\beta\gamma$ dans une gamme de concentration de 0,08 μM à 1,46 μM , sur le complexe (His6)EC/(His6)KL fixé sur le biosenseur, ont été enregistrées. Sur ces courbes, un ajustement de simili-courbes a été effectué, sur la base d'un modèle d'interaction de Langmuir afin de déterminer la constante de dissociation (Figure 47). Ce modèle d'interaction de Langmuir est un modèle d'association simple entre deux espèces : $A + B \leftrightarrow AB$.

De cet ajustement sont extraites les constantes de vitesses d'association, k_A , et de dissociation, k_D . De ces valeurs sont déterminées les constantes d'association et de dissociation. Les résultats de l'ajustement sont donnés dans le tableau 13. L'écart entre les simili-courbes de l'ajustement et les courbes expérimentales est quantifié, et la valeur de cet écart est aussi donnée dans le tableau 13.

Le complexe formé par la protéine $\beta\gamma$ et le complexe contraint EC/KL se dissocie avec une constante de 0,220 μM environ. L'ajustement est satisfaisant, l'écart entre les courbes expérimentales et l'ajustement proposé est acceptable.

II.2. Cartographie des protéines EC et KL

L'étude par RMN du complexe $\alpha\beta\gamma$ /EC/KL indique que ce complexe se meut lentement en solution, soit du fait d'une grande taille en termes de masse moléculaire, soit du fait d'une géométrie particulière comme nous l'avons vu pour le complexe EC/KL (cf. chapitre 1). Dans le cas où l'interaction entre $\alpha\beta\gamma$ et EC/KL se ferait par l'extrémité du complexe EC/KL, supposé allongé du fait de la structure en « coiled-coil », le complexe $\alpha\beta\gamma$ /EC/KL serait alors également relativement anisotrope. Ceci pourrait expliquer des mouvements en solution relativement lents. Pour tester cette hypothèse, nous avons élaboré des protéines EC et KL tronquées de leur domaine C-terminaux. En effet, la prédiction en « coiled-coil » de la séquence primaire des protéines DivIC et FtsL ne concerne pas l'intégralité de la séquence des parties extracellulaires, et permet de définir la borne d'un domaine C-terminal, au 2/3 environ de la partie extracellulaire (cf. Introduction/C.I. Figure 19). Deux nouvelles protéines ont donc été élaborées, EC* et KL* (Figure 48).

La formation d'un complexe entre les protéines EC et KL*, ainsi qu'entre les protéines EC* et KL, a été vérifiée par copurification par chromatographie d'affinité à la résine NI-

EC His6-TEV- BVSALKEKVSALKEKVSALKEKVSALKEKVS SYQQLLQRRQQLADLQYQYQLSDERDKETAFATKLKDE
EC* His6-TEV- BVSALKEKVSALKEKVSALKEKVSALKEKVS SYQQLLQRRQQLADLQYQYQLSDERDKETAFATKLKDE

KL His6- KVSALKEKVSALKEKVSALKEKVSALKEKVS MOTKLLQVNDLTKINAQIEEKKTELDDAKQEVNELRAERLKEIANS
KL* His6- KVSALKEKVSALKEKVSALKEKVSALKEKVS MOTKLLQVNDLTKINAQIEEKKTELDDAKQEVNELRA

Figure 48 : Séquence primaire des protéines EC, KL, EC* et KL*. La séquence primaire de l'étiquette polyhistidine est symbolisée par « His6 », et celle du site de clivage de la protéase Tobacco Etched Virus, par les lettres TEV.

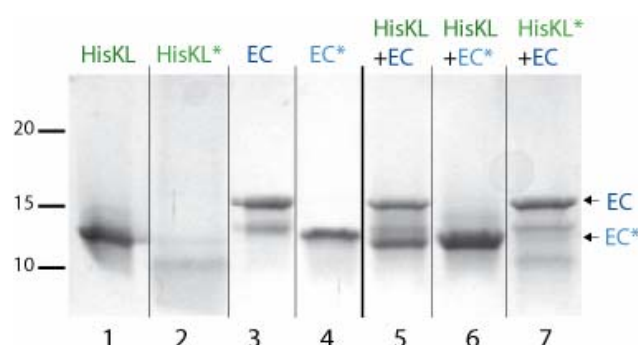


Figure 49 : Vérification de la formation des complexes EC*/KL et EC/KL*. SDS-PAGE coloré au bleu de Coomassie. Les protéines EC (piste 3) et EC* (piste 4), dont l'étiquette polyhistidine a été clivée par la protéase TEV, sont coélues avec les protéines KL (piste 1) et KL* (piste 2), d'une chromatographie d'affinité au nickel. On retrouve EC (indiquée par une flèche) dans l'élution du mélange EC/KL (piste 5) et dans celle du mélange EC/KL* (piste 7), on retrouve EC* (indiquée par une flèche) dans l'élution du mélange EC*/KL (piste 6).

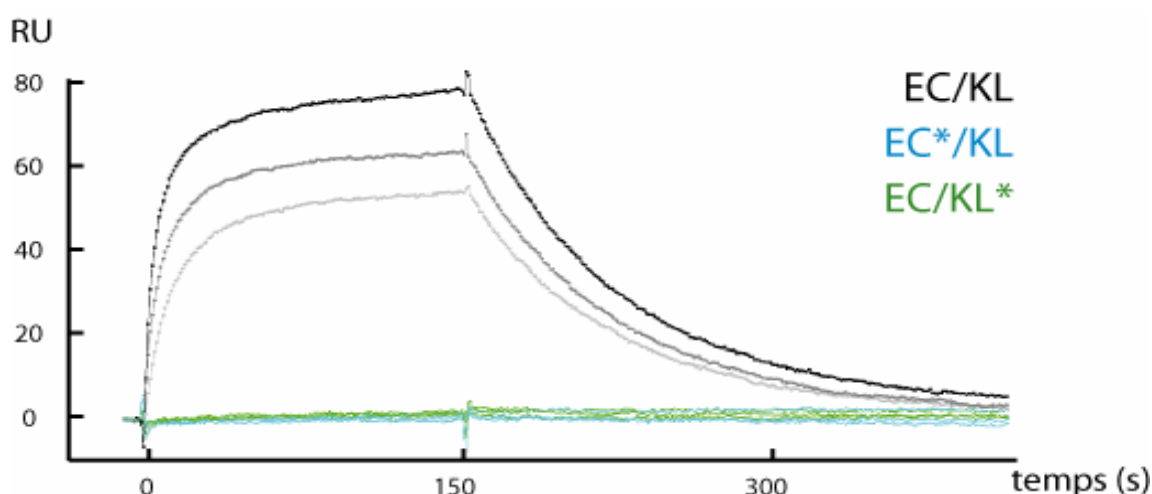


Figure 50 : Comparaison de l'interaction de la protéine β avec le complexe EC/KL, le complexe EC*/KL et le complexe EC/KL*. Les courbes d'interaction de la protéine β à différentes concentrations (1,55 μ M ; 0,77 μ M et 0,39 μ M), avec EC/KL (noir, gris foncé et gris clair, respectivement), EC*/KL (bleu) et EC/KL* (vert), fixé sur des pistes séparées d'un biosenseur, ont été suivies par résonance plasmonique de surface, avec un instrument Biacore®.

NTA®, en utilisant des protéines EC et EC*, dont l'étiquette polyhistidine avait été préalablement clivée (Figure 49).

Nous avons donc comparé l'interaction de la protéine $\beta\gamma$ avec les différents complexes contraints, EC/KL, EC*/KL et EC/KL*, par analyse d'interaction de biomolécules par résonance plasmonique de surface, avec un instrument Biacore (Figure 50). Pour ces mesures, les protéines EC ; EC*, KL et KL* portent toutes leur étiquette polyhistidine. Les complexes (His6)EC/(His6)KL ; (His6)EC*/(His6)KL et (His6)EC/(His6)KL* sont fixés jusqu'à atteindre un niveau de fixation équivalent (environ 550 RU). L'instrument Biacore utilisé (Biacore 3000) permettant l'utilisation de 4 pistes en parallèle, les trois complexes sont testés simultanément. La piste avec le complexe (His6)EC/(His6)KL a donc permis de valider le test, pour chaque essai. Sur la 4^e piste, aucun ligand n'est fixé, ce qui permet de mesurer l'interaction de l'analyte sur le biosenseur.

Il n'a pas été possible de détecter une interaction entre la protéine $\beta\gamma$ et le complexe EC*/KL, ni entre la protéine $\beta\gamma$ et le complexe EC/KL* (Figure 50).

Ces expériences ont été reproduites, de manière indépendante, trois fois. Lors de la mise au point du protocole, l'identité des protéines dans les mélanges a été vérifiée par spectrométrie de masse : deux protéines de 12705 Da et 11895 Da pour le mélange EC/KL ($M_{théo}$ EC = 12 715 Da ; $M_{théo}$ KL = 11 905 Da), deux protéines de 10527 Da et 11909 Da pour le mélange EC*/KL ($M_{théo}$ EC* = 10 660 Da ; $M_{théo}$ KL = 11 905 Da), trois protéines 12 717 Da, 10184 Da et 9209 Da pour le mélange EC/KL* ($M_{théo}$ EC = 12 715 Da ; $M_{théo}$ KL = 9 203 Da). La protéine à 10 184 Da est vraisemblablement un contaminant, ou un produit de dégradation de la protéine EC.

Il semble que les domaines C-terminaux des protéines DivIC et FtsL soient, tous les deux indispensables à l'interaction avec la protéine $\beta\gamma$. Il n'a pas été non plus possible de détecter d'interaction avec la protéine $\alpha\beta\gamma$, i.e la partie extracellulaire de DivIB entière (données non présentées), ce qui indique que l'interaction entre le complexe formé des parties extracellulaires de DivIC et de FtsL et la partie extracellulaire de DivIB se fait en effet, via les domaines C-terminaux des protéines DivIC et FtsL. Le domaine β de DivIB se fixerait donc plutôt à l'extrémité du complexe EC/KL. L'anisotropie résultant d'un complexe très allongé serait compatible avec la non résolution des pics des spectres de RMN présentés dans la section I de ce chapitre.

Echantillon	1	2	3	4	5	6	7
$[\beta\gamma]$ (μM)	0	70	240	270	300	340	340
$[\text{EC/KL}]$ (μM)	270	270	270	270	270	50	0
rapport	0 : 1	0,2 : 1	0,9 : 1	1 : 1	1 : 0,9	1 : 0,2	1 : 0
$I(0)$ ($\cdot 10^{-3}$)	201 ± 10	228 ± 10	303 ± 10	305 ± 10	320 ± 11	126 ± 9	85 ± 9
M_{app} (kDa)	$55,4 \pm 8,3$	$52,1 \pm 7,5$	$47,9 \pm 6,4$	$46,2 \pm 6,1$	$45,8 \pm 6,2$	$28,6 \pm 4,9$	$22,3 \pm 4,6$

Tableau 14 : Composition des échantillons de mélange de la protéine $\beta\gamma$ et du complexe EC/KL, et leur valeur de diffusion à l'origine. Le calcul de la masse de l'objet diffusant se fait avec la formule citée dans le texte au chapitre 1 (Eq. 36) et la valeur de diffusion incohérente du tampon H_2O . Les erreurs sur le calcul de la masse prennent en compte une erreur sur l'estimation de la concentration de 10%.

III. Modélisation du complexe trimérique par SANS

III.1. Première analyse des données

Afin d'obtenir plus d'informations structurales sur l'association entre le complexe contraint EC/KL et la protéine $\beta\gamma$, nous avons poursuivi l'étude par des expériences de diffusion aux petits angles de neutrons. La protéine $\alpha\beta\gamma$, correspondant à la partie extracellulaire de DivIB entière n'a pas été utilisée pour les raisons déjà évoquées au chapitre 2 (cf. Chap.2/II.4.4).

Plusieurs mélanges de protéines $\beta\gamma$ avec du complexe EC/KL ont été préparés pour couvrir une gamme de stoechiométrie allant d'un rapport entre $\beta\gamma$ et EC/KL de 0:1 à 1:0, en passant par une stoechiométrie 1:1. Les courbes de diffusion de ces échantillons ont été mesurées, et les valeurs d'intensité à l'origine en ont été extraites. De ces intensités à l'origine, la masse moyenne des objets diffusants a été déterminée, grâce à la calibration absolue de l'instrument (cf. Introduction/B.II.2.1). Ces valeurs sont reportées dans le tableau 14.

La masse de l'objet diffusant dans l'échantillon 1, correspondant à un échantillon de EC/KL seul, correspond bien à la masse du complexe $(EC/KL)_2$ de 50 kDa, comme attendu (cf. Chapitre 1). De même, la masse de l'objet diffusant dans l'échantillon 7, correspondant à un échantillon de $\beta\gamma$, correspond bien à la masse d'un monomère de $\beta\gamma$ de 20 kDa, comme attendu (cf. Chapitre 2). Par contre, la masse moyenne des objets diffusants des échantillons de stoechiométries intermédiaires ne correspondent ni à la somme de la masse du complexe $(EC/KL)_2$ et de la masse du monomère de $\beta\gamma$, ($50 + 20 \text{ kDa} = 70 \text{ kDa}$), ce qui serait le cas pour un complexe $\beta\gamma/(EC/KL)_2$, ni à la moyenne des espèces $\beta\gamma$ et $(EC/KL)_2$ qui n'interagiraient pas (38 kDa).

Étant donné que la masse apparente de l'objet diffusant dans l'échantillon correspondant à une stoechiométrie 1:1 (45 kDa) est très proche de la masse d'un complexe composé d'un monomère de $\beta\gamma$ et d'un complexe $(EC/KL)_1$, ($20 \text{ kDa} + 25 \text{ kDa} = 45 \text{ kDa}$), un modèle d'interaction est proposé, où le complexe se formerait entre $\beta\gamma$ et la forme $(EC/KL)_1$ du complexe contraint. L'association de $\beta\gamma$ avec EC/KL induirait donc la dissociation du complexe $(EC/KL)_2$ en 2 complexes $(EC/KL)_1$.

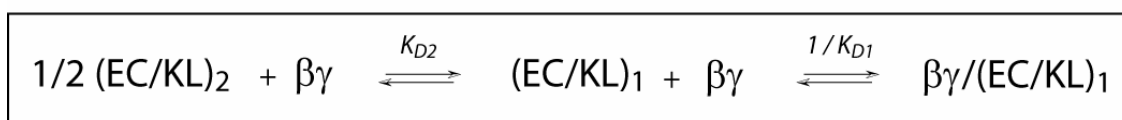


Figure 51 : Mécanismes réactionnels proposés, entre les espèces d'un mélange de complexe EC/KL et de $\beta\gamma$. L'équation chimique encadrée est la reformulation des deux équations indiquées au dessus.

III.2. Présentation d'un modèle d'association entre $\beta\gamma$ et EC/KL

Dans ce modèle d'interaction, il y aurait deux équilibres d'association/dissociation dans un échantillon d'un mélange de la protéine $\beta\gamma$ et du complexe EC/KL : l'équilibre entre le complexe $\beta\gamma/(EC/KL)_1$ et les espèces libres $\beta\gamma$ et $(EC/KL)_1$, et l'équilibre entre le complexe $(EC/KL)_2$ et le complexe $(EC/KL)_1$ (Figure 51).

La constante de dissociation, K_{D2} , du complexe $(EC/KL)_2$ a pu être estimée à partir des données expérimentales de diffusion aux petits angles de rayons X (cf. chapitre 1/V.). La valeur de cette constante serait de l'ordre de 25 μM environ. Une constante de dissociation entre $\beta\gamma$ et EC/KL a été déterminée par analyse d'interaction de biomolécules par résonance plasmonique de surface. La valeur de cette constante est d'environ 0,22 μM (cf. II.1). D'autre part, la constante de dissociation de $(EC/KL)_2$ en $(EC/KL)_1$ indique que les échantillons très dilués de EC/KL utilisés pour les expériences d'analyse d'interaction de biomolécule par résonance plasmonique de surface, sont composés majoritairement de complexe $(EC/KL)_1$. En effet, la concentration de ces échantillons est de 0,1 μM . L'estimation par le calcul des concentrations des espèces libres, en utilisant une constante K_{D2} de 25 μM , donne une composition majoritaire en $(EC/KL)_1$ (99%). Donc, la constante déterminée par analyse d'interaction de biomolécule par SPR, peut être considérée comme la constante de l'association entre la protéine $\beta\gamma$ et le complexe $(EC/KL)_1$.

Pour des échantillons équimolaires de protéine $\beta\gamma$ et de EC/KL, la composition serait alors déterminée par les équations suivantes (Eq. 39 à 42), où K_{D1} et K_{D2} sont les constantes de dissociations des réactions de complexation (Figure 51), et C_0 , la concentration totale de EC/KL ou $\beta\gamma$ (cf. Annexe).

$$\text{Eq. 39 : } \frac{2}{K_{D1}K_{D2}}[(EC/KL)_1]^3 + \left(\frac{1}{K_{D1}} + \frac{2}{K_{D2}}\right)[(EC/KL)_1]^2 + [(EC/KL)_1] = C_0$$

$$\text{Eq. 40 : } [(EC/KL)_2] = \frac{1}{K_{D2}}[(EC/KL)_1]^2$$

$$\text{Eq. 41 : } [\beta\gamma] = C_0 - [\beta\gamma/(EC/KL)_1]$$

$$\text{Eq. 42 : } [EC/KL]_{\text{total}} = [\beta\gamma]_{\text{total}} = C_0$$

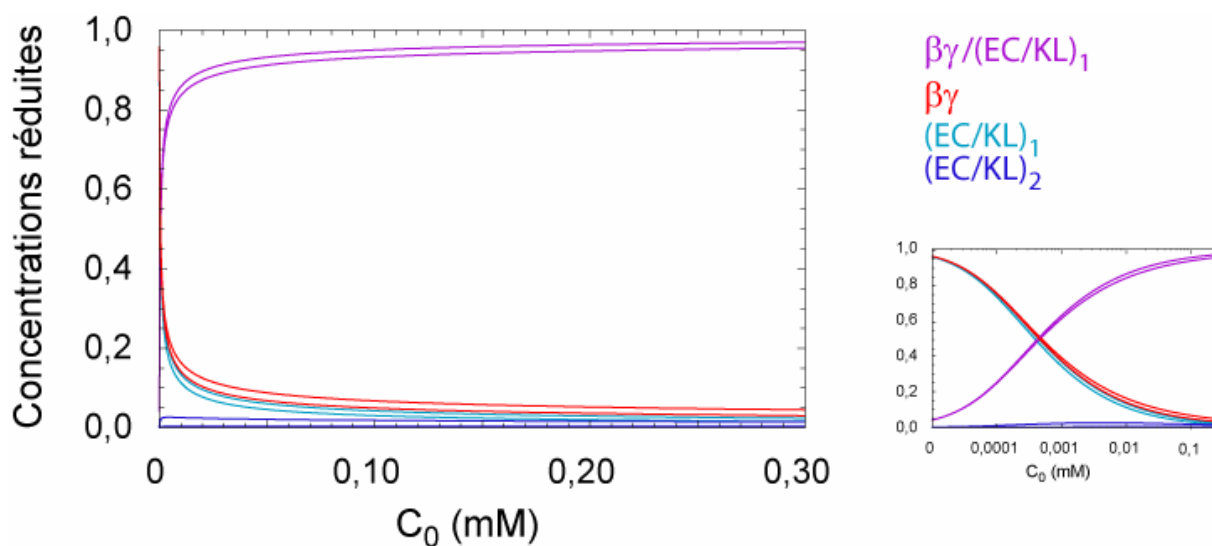


Figure 52 : Composition d'échantillons équimolaires de mélange de protéine $\beta\gamma$ et de complexe EC/KL, en fonction de la concentration, et de la constante de dissociation K_{D2} . La concentration de chaque espèce ((EC/KL)₂ en bleu foncé ; (EC/KL)₁ en bleu clair ; $\beta\gamma$ /(EC/KL)₁ en violet ; $\beta\gamma$ en rouge) est donnée par rapport à la concentration C_0 , en fonction de cette concentration C_0 , en échelle linéaire (à gauche) et logarithmique (à droite). Pour chaque espèce, deux courbes sont représentées : la concentration réduite calculée pour $K_{D2} = 5 \mu\text{M}$ et pour $K_{D2} = 50 \mu\text{M}$. Les concentrations ont été calculées avec $K_{D1} = 0,22 \mu\text{M}$.

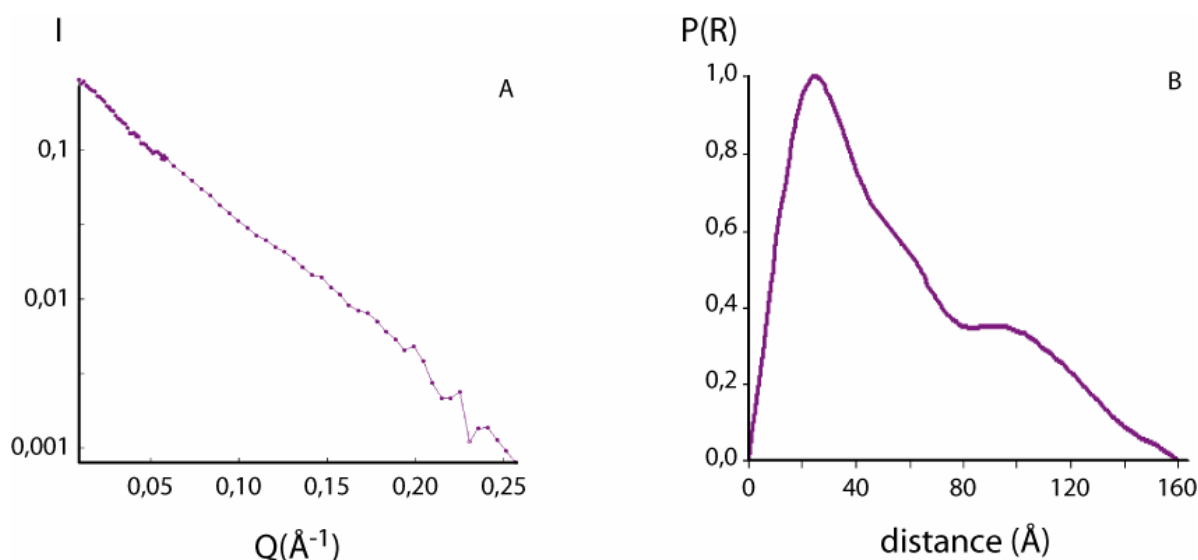


Figure 53 : Courbe de diffusion et fonction de distribution de distances de l'échantillon équimolaire de $\beta\gamma$ et de complexe EC/KL. L'intensité diffusée, en échelle logarithmique, est représentée en fonction de l'angle de diffusion (A). La distribution de distances déterminée à partir de la courbe de diffusion présentée, est normalisée par rapport à son maximum (B).

La composition de tels échantillons est dépendante de la concentration des différentes protéines. Dans le cas d'un mélange équimolaire, la composition calculée à l'aide des équations 39 à 41 ci-dessus est représentée dans la figure 52. Les calculs ont été réalisés avec une valeur de 0,22 μM pour la constante de dissociation K_{D1} (cf. II.1). Pour la constante de dissociation K_{D2} , plusieurs calculs ont été faits avec des valeurs de K_{D2} différentes, afin d'évaluer l'impact dû à une erreur d'estimation de cette constante (cf. Chapitre 1/V.).

Avec ce modèle, on remarque que, pour une concentration supérieure à 0,10 mM, les échantillons peuvent être considérés comme, majoritairement (>90%) composés de complexe $\beta\gamma/(\text{EC/KL})_1$. On remarque aussi qu'une erreur importante de l'estimation de la constante de dissociation du complexe $(\text{EC/KL})_2$, K_{D2} , estimée à partir des données de diffusion aux petits angles de rayons X, n'a que peu de conséquence. En effet, l'impact d'une variation d'un ordre de grandeur (5 à 50 μM) de cette constante sur la composition du système est négligeable.

L'échantillon équimolaire utilisé pour mesurer la courbe de diffusion d'un complexe $\beta\gamma/\text{EC/KL}$ présentée au paragraphe III.1 de ce chapitre, à la concentration de 0,27 mM, était donc très probablement composé majoritairement (94%) de complexe $\beta\gamma/(\text{EC/KL})_1$, en se basant sur ce modèle d'interaction. Ce modèle utilisé (Figure 52) rend compte des masses apparentes des échantillons (Tableau 14).

III.3. Analyse de la courbe de diffusion

La courbe de diffusion de l'échantillon équimolaire de $\beta\gamma$ et de EC/KL à 0,27 mM, peut donc être exploitée en le considérant comme composé très majoritairement de complexe $\beta\gamma/(\text{EC/KL})_1$. Deux approches d'analyses ont été utilisées, l'une par modélisation *ab initio* de la forme globale du complexe à partir de la courbe de diffusion, l'autre par ajustement de la courbe de diffusion avec une simili-courbe de diffusion calculée à partir d'un modèle par corps rigides.

III.3.1. Modélisation *ab initio*

La fonction de distribution de distances de la courbe de diffusion du complexe $\beta\gamma/(\text{EC/KL})_1$ a été déterminée avec le logiciel GNOM (Semenyuk et Svergun, 1991). La meilleure estimation de cette fonction a été trouvée pour une distance maximale de la protéine, d_{max} , de 160 Å (Figure 53).

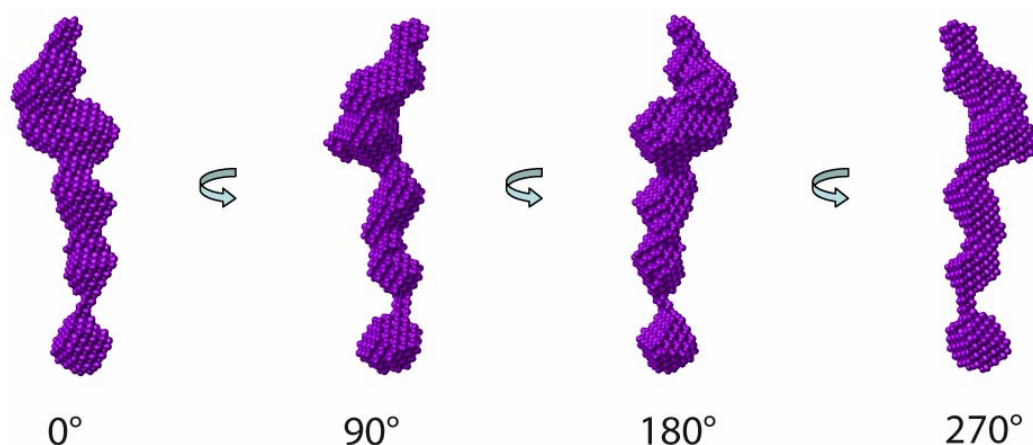


Figure 54 : Modèle *ab initio* du complexe $\beta\gamma/(EC/KL)_1$. Les différentes représentations du modèle correspondent à des rotations de 90° autour de l'axe vertical.

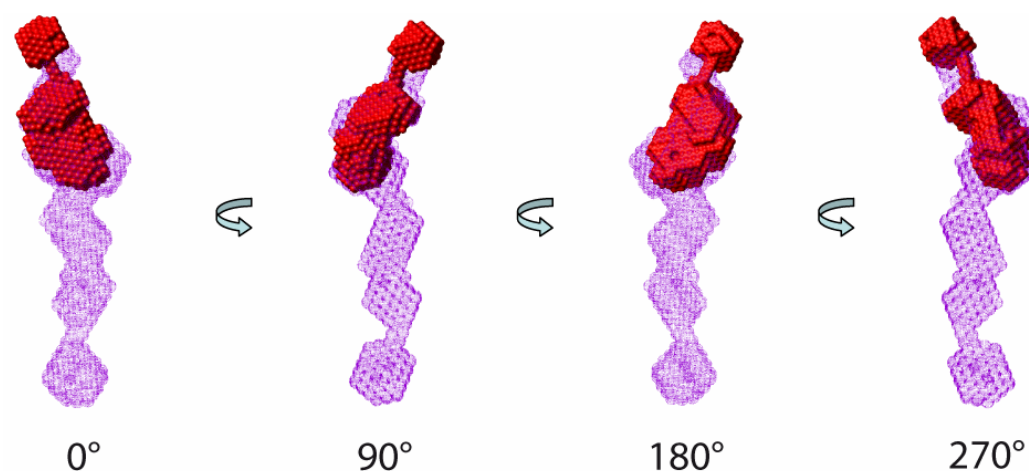


Figure 55 : Comparaison de l'enveloppe du modèle du complexe $\beta\gamma/(EC/KL)_1$ avec le modèle des domaines $\beta\gamma$, obtenus par modélisation *ab initio*. La superposition, du modèle de la protéine $\beta\gamma$ (en rouge) et du modèle du complexe $\beta\gamma/EC/KL$ (en violet) a été effectuée à la main. Les différentes représentations du modèle correspondent à des rotations de 90° autour de l'axe vertical.

Une modélisation *ab initio* de la forme du complexe a ensuite été réalisée avec le logiciel DAMMIN (Svergun, 1999), à partir de cette fonction de distribution de distances. Une première série de modélisations a été réalisée à partir d'une sphère de 160 Å de diamètre, remplie de 9688 simili-atomes de 1,9 Å de rayon. Les résultats de ces modélisations montrant systématiquement une forme allongée, une autre série de modélisation a été réalisée à partir d'un cylindre de 160 Å de long et de 40 Å de diamètre, pour augmenter la résolution des modèles. 4 modèles ont été obtenus indépendamment, et ces 4 modèles sont similaires. Un de ces modèles est présenté en figure 54, sous différentes projections.

Ces modèles représentent une forme allongée, dont la masse n'est pas uniformément répartie sur toute la longueur, mais qui présente un excédent de masse excentré. L'excédent de masse excentré rappelant la forme du modèle de $\beta\gamma$, une superposition des modèles a été effectuée manuellement (Figure 55). On remarque que le modèle de $\beta\gamma$ remplit assez bien cette zone excentrée du modèle de $\beta\gamma/(EC/KL)_1$.

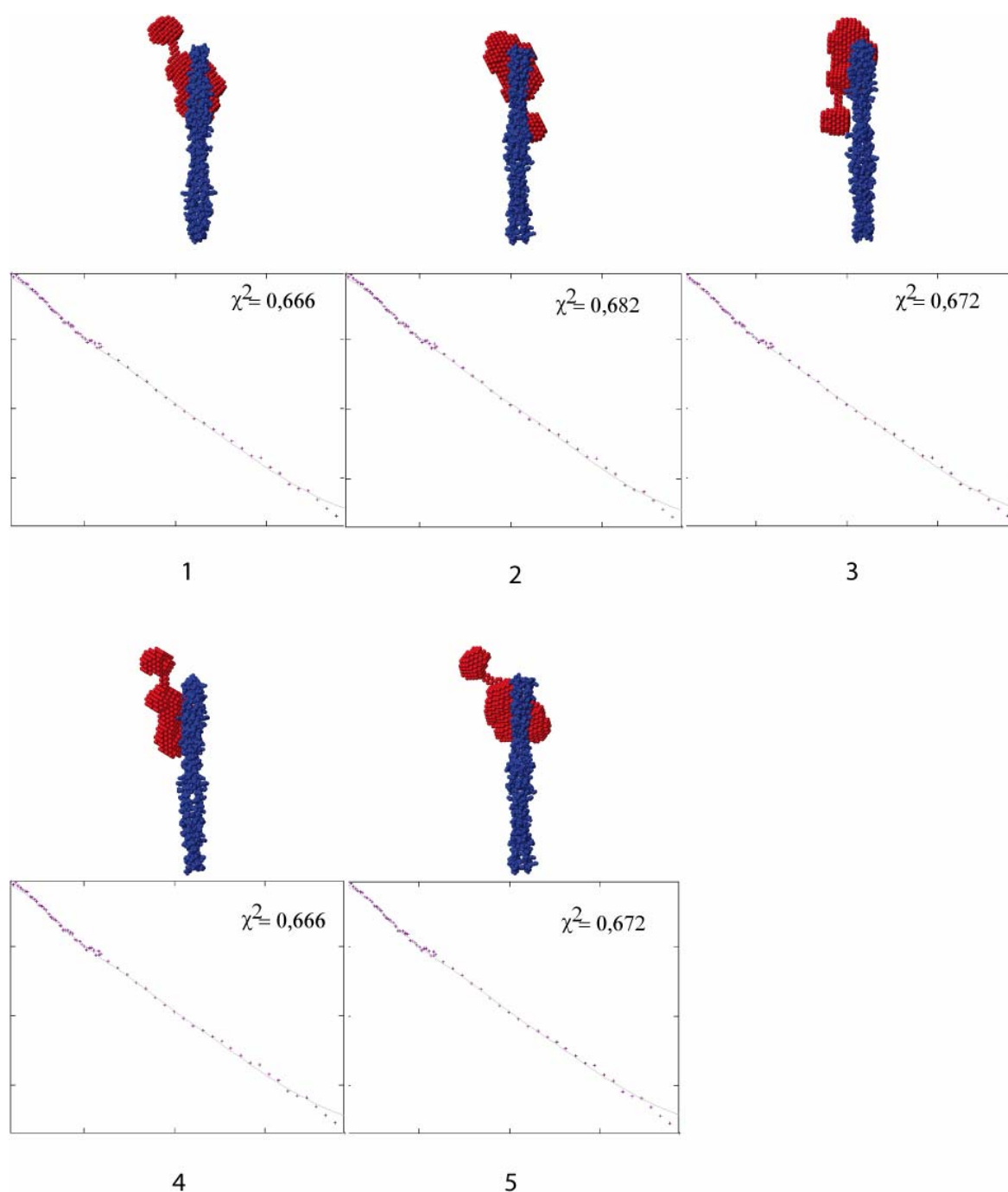


Figure 56 : **Modélisation automatique par corps rigides, de la courbe de diffusion expérimentale du complexe $\beta\gamma$ /EC/KL.** Les modèles (numérotés de 1 à 5) de complexe, ont été obtenus par ajustement de la courbe de diffusion expérimentale par la simili-courbe de diffusion calculée à partir du complexe d'un modèle de structure de $\beta\gamma$ (obtenue par modélisation *ab initio*) et d'un modèle de « coiled-coil » (code PDB : 1d7m), avec le logiciel MASSHA. Le modèle de structure de $\beta\gamma$ est coloré en rouge, la structure du « coiled-coil » en bleu. La simili-courbe finale (en gris) est comparée à la courbe expérimentale (en violet), et l'écart entre ces deux courbes est indiqué, pour chaque modèle.

III.3.2. Ajustement des corps rigides

Pour ajuster la courbe de diffusion de l'échantillon considéré comme composé de complexe $\beta\gamma/(EC/KL)_1$, nous avons utilisé le modèle à basse résolution de la protéine $\beta\gamma$ (cf. chapitre 2) comme corps rigide de la protéine $\beta\gamma$, et la structure tridimensionnelle du « coiled-coil » de la cortexilline (code PDB : 1d7m) pour $(EC/KL)_1$. En effet, une grande région de la séquence primaire de EC/KL est prédite pour former un « coiled-coil », et le modèle *ab initio* montre bien une forme allongée, compatible avec une structure en « coiled-coil » pour EC/KL. La structure « coiled-coil » utilisée (code PDB : 1d7m) comporte un nombre d'acides aminés comparable, légèrement inférieur, au complexe EC/KL : 200 acides aminés pour le modèle, alors que le complexe EC/KL est composé de 216 acides aminés. L'ajustement a été fait par comparaison à la courbe de diffusion expérimentale, d'une courbe de diffusion calculée à partir d'un modèle composé de la structure basse résolution de $\beta\gamma$ et de la structure du « coiled-coil » modèle.

L'ajustement a été réalisé de deux manières : automatique avec le logiciel MASSHA (Konarev *et al.*, 2001), et manuellement, en utilisant les logiciels MOLMOL (Koradi *et al.*, 1996) et CRYSON (Svergun, 1995).

L'ajustement informatique propose un modèle de complexe, final, après une minimisation de l'écart entre la courbe de diffusion expérimentale et la simili-courbe calculée à partir du modèle de complexe qui est réarrangé, de manière itérative.

Cet ajustement automatique donne systématiquement un positionnement de la structure de $\beta\gamma$, sur le côté d'une extrémité du « coiled-coil », quelle que soit les positions de départ de l'ajustement. Le centre de la structure de $\beta\gamma$ se retrouve systématiquement au trois quart de la longueur du « coiled-coil » (Figure 56).

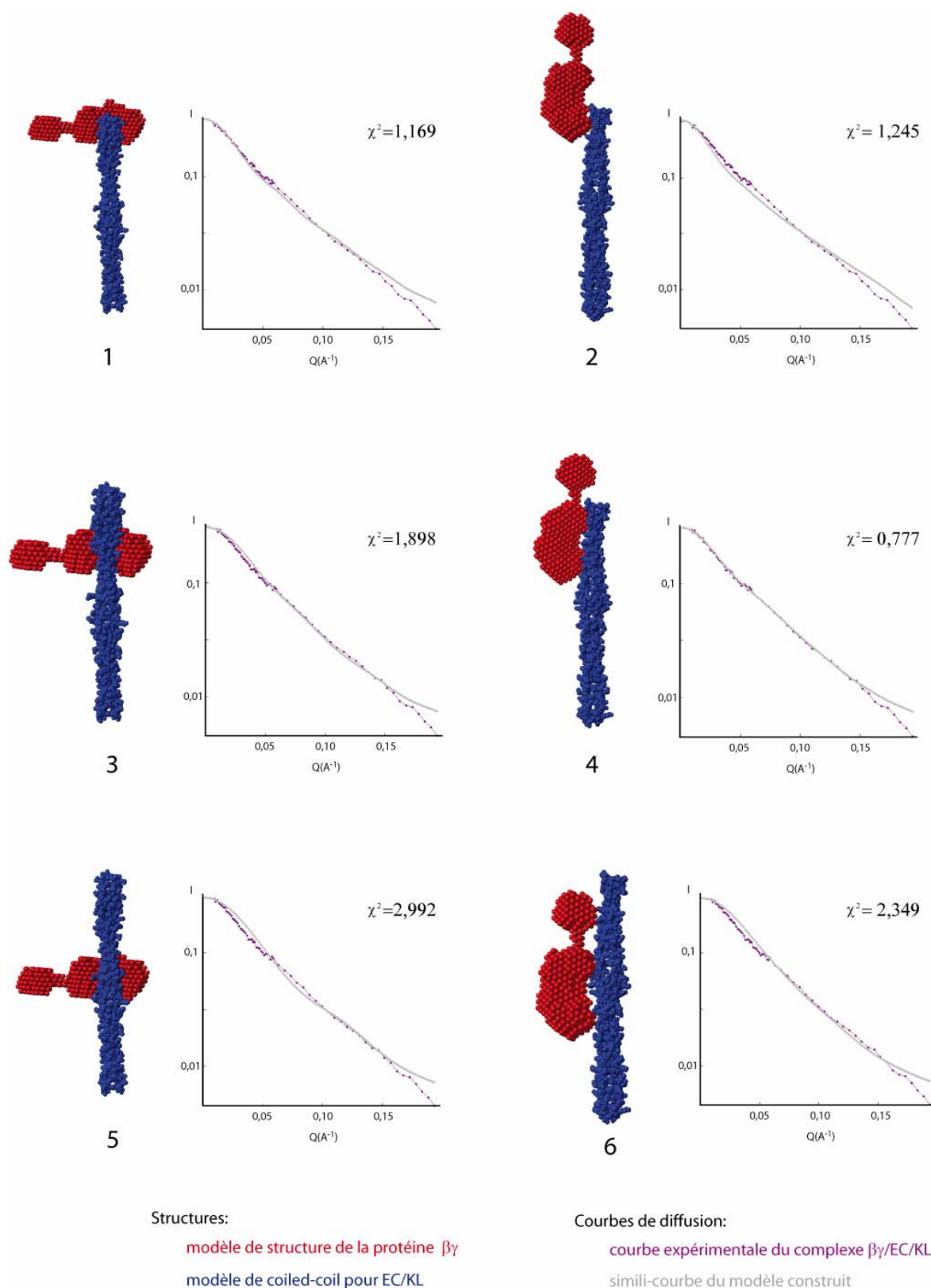


Figure 57 : **Modélisation manuelle par corps rigides, de la courbe de diffusion expérimentale du complexe $\beta\gamma$ /EC/KL.** Les modèles (numérotés de 1 à 6) de complexes ont été construits à la main, à partir du modèle de structure des doamines $\beta\gamma$ (obtenue par modélisation *ab initio*, en rouge) et d'un modèle de « coiled-coil » (code PDB : 1d7m, en bleu). Les simili-courbes de diffusion de ces modèles construits ont été calculées avec le logiciel CRYSON (en gris) et comparées à la courbe expérimentale (en violet).

Manuellement, 5 modèles ont été construits avec le logiciel MOLMOL (Koradi *et al.*, 1996). La courbe de diffusion de chaque modèle a été calculée avec le logiciel CRYSON (Svergun *et al.*, 1998) et comparée à la courbe expérimentale (Figure 57).

Les simili-courbes des modèles où le modèle de $\beta\gamma$ est au centre du « coiled-coil », parallèlement ou perpendiculairement donnent un ajustement assez éloigné de la courbe expérimentale ($\chi^2 = 2,992$ et $2,349$ pour les modèles 5 et 6). Les simili-courbes des modèles où le modèle de $\beta\gamma$ est excentré le long du « coiled-coil » sont plus proches de la courbe expérimentale. On peut d'ailleurs remarquer que le meilleur ajustement est obtenu avec la simili-courbe de diffusion du modèle 4, où le modèle de $\beta\gamma$ est au trois quart du « coiled-coil », parallèlement à lui ($\chi^2 = 0,777$).

Les deux méthodes de modélisation par corps rigides (automatique et manuelle) indiquent toutes les deux que les modèles qui donnent les meilleurs ajustements de la courbe expérimentale, sont ceux où la protéine $\beta\gamma$ est excentrée au trois quart du « coiled-coil ». Il est toutefois important de garder en tête, que le modèle de « coiled-coil » utilisé ne représente que partiellement le complexe $(EC/KL)_1$. En effet, cette structure de « coiled-coil » est longue de 100 acides aminés, alors que le « coiled-coil » du complexe EC/KL n'est prédit que sur une longueur de 72 acides aminés seulement.

Les deux approches (modélisation *ab initio* et par corps rigides) montrent que la protéine $\beta\gamma$, dans le complexe $\beta\gamma/(EC/KL)_1$, est située à un bout d'une forme allongée qui est vraisemblablement, le complexe $(EC/KL)_1$. Ce complexe est en effet anisotrope, ce qui est cohérent avec l'hypothèse expliquant les spectres de RMN de la section I de ce chapitre. Le positionnement de la protéine $\beta\gamma$ à un bout du complexe EC/KL est aussi cohérent avec les expériences d'analyse d'interaction de biomolécules, qui montraient que la protéine $\beta\gamma$ interagit avec les régions C-terminales des protéines EC et KL.

Chapitre 4. Matériels et méthodes

I. Constructions génétiques

I.1. Le complexe contraint EC/KL

La construction génétique codant pour la protéine EC en fusion avec une étiquette Strep (Step-e5-domaine extracellulaire de DivIC) a été introduite dans un vecteur pET 30b (Novagen) par André Zapun.

La construction génétique codant pour la protéine EC en fusion avec une étiquette polyhistidine (His6-TEV-e5-domaine extracellulaire de DivIC) a été introduite dans le vecteur pET 30b (Novagen) par André Zapun.

La construction génétique codant pour la protéine initiale KL (His6-k5-domaine extracellulaire de FtsL) a été introduit dans le site de clonage 2 d'un vecteur pET Duet (Novagen) par Marjolaine Noirclerc-Savoye (Noirclerc-Savoye *et al.*, 2005).

La construction génétique codant pour la protéine KL utilisée dans ce travail de thèse (excepté pour les premiers spectres de RMN) a été obtenue par insertion d'un codon TGG (codant pour le tryptophane) entre les séquences de l'étiquette polyhistidine et celle du premier motif K. Ce codon a été inséré par mutagenèse dirigée avec les oligomères GCATCACCATCACCATCACTGGGGTGGCAAGGTATCCGC et GCGGATACCTTGCC ACCCCAGTGATGGTGGTGGTGGTGGTGC par le système QuickChange® II KL Site-Directed Mutagenesis (Stratagene). La protéine synthétisée est His6-W-k5-domaine extracellulaire de FtsL.

Le plasmide pour l'expression de la protéine KL (^RKana ; ^SAmp) en milieu deutéré, a été obtenu par insertion du gène de résistance à la kanamycine dans le gène de résistance à l'ampicilline par le système de EZ-T5TM de Epicentre. Le plasmide a été obtenu après sélection d'une souche de bactéries BL21 DE3 transformée avec le produit de transposition poussant sur une gélose de milieu LB contenant de la kanamycine à 30 mg/L et ne poussant pas sur une gélose de milieu LB contenant 100 mg/L d'ampicilline. L'insertion d'un seul transposon a été vérifiée par digestion du plasmide avec l'enzyme de restriction *EcoRI* et analyse des tailles des fragments par gel d'agarose.

Le plasmide de co-expression des protéines EC et KL a été obtenu par insertion du gène codant pour la protéine EC (extrait du plasmide pET 30b par digestion par les enzymes

*Xba*I et *Xho*I) entre les sites de restriction *Xba*I et *Sal*I du plasmide pET Duet + KL. L'insertion du gène a ensuite été vérifiée par séquençage.

La construction génétique codant pour la protéine KL* a été obtenue par insertion d'un codon TAG (codant pour l'arrêt de transcription) dans la séquence codant pour la protéine KL. Ce codon a été inséré par mutagenèse dirigée avec les oligomères CAATGAACTATTACGTGCATAGTATACGAAAGAAATTGC et GCAATTTCTTTCGTATACTATGCACGTAATAGTTCATTG, par le système QuickChange® II KL site-Directed Mutagenesis (Stratagene). La protéine synthétisée est His6-W-k5-domaine extracellulaire de FtsL prédit en coiled-coil (de la méthionine 44 à l'alanine 82).

La construction génétique codant pour la protéine EC* a été introduite dans le vecteur pLIM09 (LIM/IBS), par la plateforme RoBioMol, pour permettre l'expression en fusion avec une étiquette polyhistidine et un site de clivage de la protéase Tobacco Etched Virus (TEV). Le produit transcrit est MAHHHHHHGH HHQ –TEV- e5-domaine extracellulaire de DivIC prédit en coiled-coil (de la sérine 55 au glutamate 93).

I.2. La partie extracellulaire de DivIB

Toutes les constructions permettant l'expression du domaine extracellulaire entier ou tronqué de DivIB ont été obtenues par amplification par PCR à partir du chromosome de la souche R6 de *S. pneumoniae* et introduction dans des vecteurs d'expression.

Le fragment de gène codant pour le domaine extracellulaire de DivIB (IB ou $\alpha\beta\gamma$), a été introduit dans le vecteur pGex4T1 (GE Healthcare) par Marjolaine Noirclerc-Savoye, pour permettre l'expression en fusion avec la protéine glutathion S-transférase (GST) et un site de clivage de la protéase Tobacco Etched Virus (TEV). Le produit transcrit est GST-TEV- $\alpha\beta\gamma$.

Les fragments de gène codant pour les protéines α , β , et $\alpha\beta$ ont été introduits dans le vecteur pLX06 (Protein'EXpert), par la plateforme RoBioMol, pour permettre l'expression en fusion avec la protéine glutathion S-Transférase (GST) et un site de clivage de la protéase Tobacco Etched Virus (TEV). Les produits transcrits sont GST-TEV-GHHHQL- α ; GST-TEV-GHHHQL- β et GST-TEV-GHHHQL- $\alpha\beta$.

Le fragment de gène codant pour la protéine $\beta\gamma$ a été introduit dans le vecteur pGex4T1 (GE Healthcare) par Isabelle Petit-Haertlein et moi-même. L'amplification du gène a été réalisée avec les primers AGGATCCGAGAACCTGTACTTCCAGGGCGAATATG ATATTGTGGCCTACTATATTTCTGG et AGAATTCCTAGCGACGCGATGAACGCTG.

Un site de clivage pour la protéase TEV est encodé dans le premier primer (en gras). Le produit de PCR a été inséré entre les sites de restriction *Bam*HI et *Eco*RI en fusion avec le gène codant pour la protéine glutathion S-Transférase (GST). Le produit transcrit est GST-TEV- $\beta\gamma$.

II. Production des protéines

Pour produire ces protéines en milieu H₂O, les bactéries *E. coli* BL21 DE3 Codon + RIL compétentes ont été transformées avec les plasmides appropriés par choc thermique à 42°C pendant 45 secondes puis cultivées dans du milieu SOC, puis sur une gélose de milieu LB en présence des antibiotiques appropriés. Pour les protéines $\alpha\beta\gamma$, $\beta\gamma$ et β , les transformations de bactéries étaient réalisées immédiatement avant la production.

II.1. Production classique

Pour produire la protéine EC, 2 L de milieu LB ont été inoculés au 1/25 avec une culture sur la nuit. Les bactéries ont été cultivées à 37°C sous agitation en présence de kanamycine (30 mg/L). Lorsque la densité optique à 600 nm du milieu a eu atteint une valeur de 0,9 environ, la transcription de la protéine sous le contrôle du promoteur *T7lac* a été induite par 1 mM d'IPTG pendant 3 heures à 37°C sous agitation.

Pour produire la protéine EC*, 1L de milieu LB a été inoculé au 1/25 avec une culture sur la nuit. Les bactéries ont été cultivées à 37°C sous agitation en présence de kanamycine (30 mg/L). Lorsque la densité optique à 600 nm du milieu a eu atteint une valeur de 1 environ, la transcription de la protéine sous le contrôle du promoteur *T7lac* a été induite par 0,5 mM d'IPTG pendant 3 heures à 37°C sous agitation.

Pour produire la protéine KL, 2 L de milieu d'auto-induction (yeast extract 5 g/L ; tryptone 10 g/L ; MgSO₄ 1mM ; (NH₄)₂SO₄ 3,3 g/L ; KH₂PO₄ 6,8 g/L ; Na₂HPO₄ 7,1 g/L ; FeCl₃ 50 μ M ; CaCl₂ 20 μ M ; MnCl₂ 10 μ M ; ZnSO₄ 10 μ M ; CoCl₂ 2 μ M ; CuCl₂ 2 μ M ; NiCl₂ 2 μ M ; Na₂MoO₄ 2 μ M ; H₃BO₃ 2 μ M ; glycérol 5 g/L ; D(+)-glucose 0,5 g/L ; lactose 5 g/L) ont été inoculés au 1/25 avec une culture sur la nuit et cultivés pendant 24 heures à 37°C sous agitation en présence d'ampicilline (100 mg/L).

Pour produire la protéine KL*, 2 L de milieu TB ont été inoculés au 1/25 avec une culture sur la nuit. Les bactéries ont été cultivées à 37°C sous agitation en présence d'ampicilline (100 mg/L). Lorsque la densité optique à 600 nm du milieu a eu atteint une

valeur de 2 environ, la transcription de la protéine sous le contrôle du promoteur *T7lac* a été induite par 0,5 mM d'IPTG pendant 3 heures à 37°C sous agitation.

Pour produire la protéine IB, ou $\alpha\beta\gamma$, 2L de milieu d'auto-induction (yeast extract 5 g/L ; tryptone 10 g/L ; MgSO_4 1mM ; $(\text{NH}_4)_2\text{SO}_4$ 3,3 g/L ; KH_2PO_4 6,8 g/L ; Na_2HPO_4 7,1 g/L ; FeCl_3 50 μM ; CaCl_2 20 μM ; MnCl_2 10 μM ; ZnSO_4 10 μM ; CoCl_2 2 μM ; CuCl_2 2 μM ; NiCl_2 μM ; Na_2MoO_4 2 μM ; H_3BO_3 2 μM ; glycérol 5 g/L ; D(+)-glucose 0,5 g/L ; lactose 5 g/L) ont été inoculés au 1/25 avec une culture sur la nuit et cultivés pendant 24 heures à 37°C sous agitation en présence d'ampicilline (100 mg/L).

Pour produire la protéine $\beta\gamma$, 1 L de milieu TB a été inoculé au 1/25 avec une culture sur la nuit. Les bactéries ont été cultivées à 37°C sous agitation en présence d'ampicilline (100 mg/L). Lorsque la densité optique à 600 nm du milieu a eu atteint une valeur de 2 environ, la transcription de la protéine sous le contrôle du promoteur *T7lac* a été induite par 0,5 mM d'IPTG pendant 3 heures à 37°C sous agitation.

Pour réaliser les tests d'expression et de purification des protéines α , β , $\alpha\beta$, et $\beta\gamma$, pour chaque protéine, plus la protéine $\alpha\beta\gamma$ (ou IB), en contrôle, 500 mL de milieu TB ont été inoculés au 1/30 avec une culture sur la nuit puis cultivés à 37°C sous agitation en présence d'ampicilline (100 mg/L). La transcription des protéines a été induite par 1 mM IPTG, lorsque le milieu eut atteint une densité optique à 600 nm de 2 environ, pendant 3 heures à 37°C sous agitation.

II.2. Production pour marquage isotopique

II.2.1. Marquage à l'isotope 15 de l'azote

Pour produire la protéine KL qui a été utilisée pour enregistrer le spectre HSQC [^{15}N ; ^1H], les bactéries ont été cultivées dans un milieu minimal, où la source d'azote est contrôlée. Pour éviter la dilution isotopique, les bactéries d'une pré-culture de 25 mL de milieu LB ont été récoltées par centrifugation (3000 g, 10 minutes, 4°C) puis re-suspendues dans 1 L de milieu minimun ^{15}N (KH_2PO_4 3g/L ; Na_2HPO_4 6g/L ; NaCl 0,5 g/L ; MgSO_4 1mM ; FeCl_3 50 μM ; CaCl_2 20 μM ; MnCl_2 10 μM ; ZnSO_4 10 μM ; CoCl_2 2 μM ; CuCl_2 2 μM ; NiCl_2 μM ; Na_2MoO_4 2 μM ; H_3BO_3 2 μM ; biotine 1 μM ; niacinamide 5 μM ; pyridoxine HCl 5 μM ; thiamine HCl 5 μM ; acide folique 1 μM ; riboflavine 1 μM ; glucose 4g/L ; $^{15}\text{NH}_4\text{Cl}$ 1 g/L). Les bactéries ont été cultivées à 37°C sous agitation en présence d'ampicilline (100 mg/L). Lorsque la densité optique à 600 nm du milieu a eu atteint 0,7 , la transcription de la protéine,

sous contrôle du promoteur *T7lac*, a ensuite été induite avec 1 mM d'IPTG pendant 3 heures à 37 °C sous agitation.

Pour produire l'échantillon de $\alpha\beta\gamma$ (ou IB), qui a été utilisé pour enregistrer le spectre HSQC [^{15}N ; ^1H], les bactéries ont été cultivées dans un milieu minimal, où la source d'azote est contrôlée. Pour éviter la dilution isotopique, l'ensemencement a été réalisé en deux étapes. Dans un premier temps, 40 mL de milieu minimum ^{15}N (KH_2PO_4 3g/L ; Na_2HPO_4 6g/L ; NaCl 0,5 g/L ; MgSO_4 1 mM ; FeCl_3 50 μM ; CaCl_2 20 μM ; MnCl_2 10 μM ; ZnSO_4 10 μM ; CoCl_2 2 μM ; CuCl_2 2 μM ; NiCl_2 μM ; Na_2MoO_4 2 μM ; H_3BO_3 2 μM ; biotine 1 μM ; niacinamide 5 μM ; pyridoxine HCl 5 μM ; thiamine HCl 5 μM ; acide folique 1 μM ; riboflavine 1 μM ; glucose 4g/L ; $^{15}\text{NH}_4\text{Cl}$ 1 g/L) ont été inoculés au centième et cultivés 15 heures à 37°C sous agitation en présence d'ampicilline (100 mg/L). Puis, 1 L de milieu minimum ^{15}N a été inoculé avec cette préculture. Après 15 heures de culture à 28°C et agitation à 550 rpm en bio-réacteur minifors (Infors) en présence d'ampicilline (100 mg/L), 500 mL de milieu minimum ^{15}N ont été ajoutés à ce litre de culture. La transcription de la protéine, sous contrôle du promoteur *Ptac*, a ensuite été induite avec 1 mM d'IPTG pendant 3 heures à 30 °C et agitation à 700 rpm.

II.2.2. Marquage à l'isotope 2 de l'hydrogène, le deutérium D

Pour produire les protéines EC et KL en milieu D_2O , les bactéries *E. coli* BL21 DE3 compétentes ont été transformées avec les plasmides pET Duet ($^{\text{R}}$ Kana) + KL pour la protéine KL et pET30b ($^{\text{R}}$ Kana) + EC pour la protéine EC, par choc thermique à 42°C pendant 45 secondes puis cultivées dans du milieu SOC, puis sur une gélose de milieu LB H_2O en présence de kanamycine (30 mg/L). Dans un premier temps, les bactéries ont été cultivées en milieu minimum Enfors H_2O ($(\text{NH}_4)_2\text{SO}_4$ 6,86 g/L ; KH_2PO_4 1,56 g/L ; Na_2HPO_4 6,48 g/L ; $(\text{NH}_4)_2\text{-Citrate}$ 0,49 g/L ; CaCl_2 20 μM ; FeCl_3 50 μM ; ZnSO_4 10 μM ; CuSO_4 2 μM ; MnSO_4 10 μM ; CoCl_2 2 μM ; MgSO_4 1 mM ; glycérol 5 g/L) à 37°C sous agitation, pour être ensuite progressivement adaptées à l'eau D_2O . De cette pré-culture en milieu minimum H_2O , elles ont été cultivées sur gélose de milieu Enfors D_2O à 37°C en présence de kanamycine (30 mg/L). Les colonies de bactéries ayant poussées sur cette gélose ont ensuite été cultivées en milieu liquide Enfors 85% D_2O /15% H_2O à 37°C sous agitation, pendant plusieurs jours, en renouvelant le milieu, pour assurer une bonne adaptation des bactéries au milieu D_2O . Les bactéries ont ensuite été cultivées au bioréacteur dans 3 L de milieu Enfors selon le protocole de production du laboratoire D-Lab (ILL) avec Martine Moulin (D-Lab/ILL) et Michael Haertlein (D-Lab/ILL).

II.3. Production en bioréacteur

Pour obtenir une plus grande quantité de bactéries, un protocole d'expression en bioréacteur a été élaboré par Benoit Gallet (LIM/IBS) et moi-même. Les bactéries d'une pré-culture de 200 mL de milieu TB cultivée à 37°C sous agitation pendant 15 heures en présence d'antibiotique ont été récoltées par centrifugation (3000 g, 15 minutes) et resuspendues dans 1 L de milieu TB, dans un bioréacteur Minifors (Infors). Les bactéries ont été cultivées 5 heures à 37°C et 700 rpm, avec une arrivée d'air dans le milieu et en présence d'antibiotique. La transcription a ensuite été induite par l'ajout de 1 mM d'IPTG dans 0,5 L de milieu TB supplémentaire, pendant 3 heures à 37°C et 1000 rpm.

Pour toutes les productions, les bactéries ont été récoltées par centrifugation (4000 g pendant 20 minutes à 4°C) et re-suspendues dans une solution de volume d'1/20, tamponnée à pH 8 par 20 mM Tris, contenant 500 mM de NaCl. Les bactéries ont ensuite été congelées à -20°C.

III. Purification

III.1. Purification du complexe contraint EC/KL et ses variants

III.1.1 Purification de (Strep)EC

Pour purifier la protéine EC en fusion avec une étiquette Strep, les bactéries ont été décongelées puis lysées par sonication aux ultrasons (Misonix) (4 minutes au total, 2 secondes toutes les 10 secondes, dans un bain de glace). Le surnageant du lysat bactérien (40 000 g pendant 20 minutes à 4°C) a été injectés à 2 mL/min sur 5 mL de résine Streptactin (IBA) à 4°C. La résine a été lavée avec 50 mL de Tris 20 mM pH 8, NaCl 150 mM. La protéine a ensuite été éluée par 2,5 mM de desthiobiotine dans 20 mM Tris pH 8, 150 mM NaCl. Les fractions d'élution ont été analysées par SDS-PAGE 15%. Les fractions intéressantes ont ensuite été réunies et conservées à 4°C.

III.1.2 Purification de (his6)EC, KL, EC* et KL*

Pour purifier ces protéines, les bactéries ont été décongelées puis lysées par sonication aux ultrasons (Misonix) (4 minutes au total, 2 secondes toutes les 10 secondes, dans un bain de glace). Le surnageant du lysat bactérien (40 000 g pendant 20 minutes à 4°C) a été injectés sur 5 mL de résine NiNTA (Qiagen) à 2 mL/min à température ambiante. La résine a été lavée

avec du Tris 20 mM pH 8, NaCl 500 mM, imidazole 20 mM. Les protéines ont ensuite été éluées par un gradient de concentration d'imidazole de 20 à 500 mM, en 50mL dans 20 mM Tris pH 8, 500 mM NaCl. Les fractions d'élution ont été analysées par SDS-PAGE 15%. Les fractions intéressantes ont ensuite été réunies et conservées à 4°C.

III.1.3. Préparation de EC/KL, EC*/KL et EC/KL*

Pour préparer ces complexes, à partir des protéines purifiées séparément, la concentration des 2 protéines du complexe a été estimée par leur absorbance à 280 nm en utilisant leur coefficient d'extinction molaire $\epsilon_{280} = 17\,420\text{ cm}^{-1}.\text{M}^{-1}$ pour (Strep)EC, $\epsilon_{280} = 5\,500\text{ cm}^{-1}.\text{M}^{-1}$ pour KL et KL*, $\epsilon_{280} = 13\,410\text{ cm}^{-1}.\text{M}^{-1}$ pour (His6)EC et $\epsilon_{280} = 4470\text{ cm}^{-1}.\text{M}^{-1}$ pour EC*. Les protéines ont ensuite été mélangées en quantité stoechiométrique. Puis, une masse d'urée a été rajoutée au mélange, de manière à obtenir une concentration finale en urée de 8 M. Le mélange a ensuite été dialysé 3 fois contre 40 fois son volume de Tris 20 mM pH 8, NaCl 150 mM à 4°C, puis conservé à 4°C.

III.1.4. Co-purification de EC/KL

Pour purifier le complexe EC/KL, soit les bactéries ayant exprimées les deux protéines ont été décongelées puis lysées (Misonix) (4 minutes au total, 2 secondes toutes les 10 secondes, dans un bain de glace) et le surnageant du lysat bactérien (40 000 g pendant 20 minutes à 4°C), soit le mélange de EC et de KL (cf. III.1.3), a été injecté sur 5 mL de résine NiNTA (Qiagen) à 2 mL/min à température ambiante. La résine a été lavée avec du Tris 20 mM pH 8, NaCl 150 mM, imidazole 20 mM. Le complexe a ensuite été élué par un gradient d'imidazole de 20 à 500 mM en 50 mL dans 20 mM Tris pH 8, 150 mM NaCl. Les fractions d'élution ont été analysées par SDS-PAGE 15%. Les fractions intéressantes ont ensuite été réunies et injectées sur 5 mL de résine Streptactin (IBA) à 2 mL/min à 4°C. La résine a été lavée avec 50 mL de Tris 20 mM pH 8, NaCl 150 mM. Le complexe a ensuite été élué par 2,5 mM de desthiobiotine dans 20 mM Tris pH8, 150 mM NaCl. Les fractions d'élution ont été analysées par SDS-PAGE 15%. Les fractions intéressantes ont ensuite été réunies et conservées à 4°C.

Le mélange a ensuite été concentré avec une unité de concentration Amicon (Millipore), jusqu'à un volume inférieur à 5 mL puis injecté sur 120 mL de résine S200 (16x60 mm) à 1 mL/min à température ambiante, préalablement équilibrée avec 20 mM Tris pH 8, 150 mM NaCl, excepté pour les spectres RMN, où l'équilibration avait été faite avec 20 mM phosphate pH 6, 150 mM NaCl. Les fractions de l'élution ont été analysées par SDS-

PAGE 15%. Les fractions intéressantes ont ensuite été réunies et concentrées avec une unité de concentration Amicon (Millipore). La concentration des échantillons a été calculée à partir de l'absorbance à 280 nm, avec le coefficient molaire $\epsilon_{280} = 22920 \text{ cm}^{-1} \cdot \text{M}^{-1}$ déterminé par une analyse en acides aminés.

Pour les échantillons partiellement deutérés, le transfert en tampon 41% D₂O/59% H₂O a été fait par une succession de dilution par 15 dans le tampon final, et de concentration jusqu'à un volume de 200 μL . La composition finale du tampon des échantillons a été vérifiée par la valeur de la transmission de neutrons ($T=0,8$).

III.2. Purification du domaine extracellulaire de DivIB et ses variants

III.2.1 Purification de $\alpha\beta\gamma$ (ou IB), $\beta\gamma$, β , par chromatographie d'affinité

Pour purifier ces protéines, les bactéries ont été décongelées puis lysées par sonication aux ultrasons (Misonix) (4 minutes au total, 2 secondes toutes les 10 secondes, dans un bain de glace). Le surnageant du lysat bactérien après centrifugation (40 000 g pendant 20 minutes à 4°C) a été mis en contact avec 10 mL de résine glutathion Sepharose (GE Healthcare) pendant 20 minutes à température ambiante en agitation douce. La résine a été lavée avec 100 mL de Tris 20 mM pH 8, NaCl 500 mM. La fusion GST a ensuite été clivée, sur la résine, par la protéase TEV pendant 2 heures à 37°C. Le produit de clivage a été récupéré par 5 lavages de la résine avec 10 mL Tris 20 mM pH 8, NaCl 150 mM. La protéase TEV a ensuite été retirée du mélange par passage sur 1 mL de résine NiNTA (Qiagen) (la protéase contient une étiquette hexahistidine). Les fractions d'élutions ont été analysées par SDS-PAGE 12,5%. Les fractions intéressantes ont ensuite été réunies et conservées à 4°C.

III.2.2. Préparation d'échantillon de protéine β par protéolyse limitée de protéine $\alpha\beta\gamma$ (ou IB)

L'échantillon de β utilisé pour enregistrer le spectre HSQC [^{15}N ; ^1H] a été préparé par protéolyse d'une partie de la préparation de IB (ou $\alpha\beta\gamma$) ^{15}N produit comme décrit aux paragraphes II.2.1 et III.2.1. Environ 200 μg de trypsine (Sigma T4549) ont été ajoutés à la protéine IB (ou $\alpha\beta\gamma$). Le mélange a ensuite été incubé pendant 20 minutes à 37°C, puis dilué 10 fois avec une solution tamponnée avec 20 mM Tris pH 8, pour diminuer la concentration de NaCl jusqu'à 15 mM environ. Le mélange a ensuite été purifié par chromatographie échangeuse d'ions. Le mélange a été injecté sur 6 mL de résine Resource Q (GE Healthcare) à

2 mL/min à température ambiante. La résine a été lavée avec 10 mL de tampon Tris 20 mM pH 8, NaCl 15 mM. Puis la protéine a été éluée avec un gradient de concentration de NaCl, de 15 à 500 mM en 50 mL. Les fractions ont été analysées par SDS-PAGE 15%. Les fractions intéressantes ont ensuite été réunies et conservées à 4°C.

III.2.3. Tests de purification des protéines α , β , $\alpha\beta$ et $\beta\gamma$

Pour estimer la quantité de protéine purifiable, les 5 lots de bactéries (α , β , $\alpha\beta$, $\beta\gamma$ et $\alpha\beta\gamma$) produit comme décrit au paragraphe II.1, ont été décongelés, puis lysés par sonication aux ultrasons (Misonix) (3 minutes au total, 2 secondes toutes les 10 secondes, dans un bain de glace). Pour chaque protéine, le surnageant du lysat bactérien après centrifugation (40 000 g pendant 20 minutes à 4°C) a été chargé sur 1 mL de résine glutathion Sepharose à 2 mL/min à température ambiante. Après lavage avec du tampon Tris 20 mM pH 8, NaCl 500 mM, la protéine a été éluée avec 10 mM de glutathion réduit dans du tampon Tris 20 mM pH 8, NaCl 500 mM. Les fractions ont ensuite été analysées par SDS-PAGE 15%. Les fractions intéressantes ont ensuite été réunies. Une mesure d'absorbance à 280 nm a permis d'estimer la quantité de protéine exprimée et purifiée. La protéine GST a été clivée par ajout de protéase TEV. Après 15 heures à 4°C sous agitation douce, les échantillons ont été centrifugés (3000 g, 20 min, 4°C) pour séparer le surnageant d'un précipité. Le surnageant et le précipité de chaque mélange, ont ensuite été analysés par SDS-PAGE 15%. Le surnageant de chaque mélange a été passé sur 1 mL de résine NiNTA (GE Healthcare) pour retirer la protéase TEV, puis dialysé contre cent volumes de tampon Tris 20 mM pH 8, NaCl 500 mM, pour retirer le glutathion réduit. Les échantillons ont ensuite été passés sur 1 mL de résine glutathion sepharose (GE Healthcare) pour retirer le domaine GST, puis concentrés avec des unités de concentration Amicon (Millipore) jusqu'à un volume de 500 μ L, avant analyse par SDS-PAGE 15%.

III.2.4. Purification par chromatographie échangeuse d'ions

Pour obtenir des échantillons de protéine $\alpha\beta\gamma$ (ou IB), $\beta\gamma$, β , avec un niveau de pureté plus élevé, une deuxième étape de purification a été réalisée. La première étape de purification de cet échantillon a été réalisée comme décrite au paragraphe précédent. L'éluion de cette chromatographie d'affinité a ensuite été dialysée afin de diminuer la concentration en NaCl jusqu'à 15 mM environ. La dialyse a été réalisée à 4°C, à travers une membrane Spectrum, contre un volume adapté de Tris 20 mM pH 8. La solution a ensuite été injectée sur 6 mL de

résine Resource Q (GE Healthcare) à 2 mL/min à température ambiante. La résine a été lavée par 10 mL de Tris 20 mM pH 8, NaCl 15 mM. Les protéines ont alors été éluées par un gradient de concentration de NaCl, de 15 à 500 mM en 50 mL. Les fractions ont été analysées par SDS-PAGE 12,5%. Les fractions intéressantes ont ensuite été réunies et conservées à 4°C.

III.2.5. Purification par chromatographie d'exclusion de taille

L'homogénéité des échantillons étant essentielles pour certaines expériences (notamment les expériences de diffusion aux petits angles), une troisième étape de purification par chromatographie par exclusion de taille a été réalisée pour préparer certains échantillons. Les protéines β , $\beta\gamma$ et $\alpha\beta\gamma$ (ou IB) ont d'abord été purifiées par chromatographie d'affinité comme décrit au paragraphe III.2.4, puis par chromatographie échangeuse d'ions comme décrit au paragraphe III.2.3. Les fractions d'éluion ont ensuite été concentrées avec une unité de concentration Amicon (Millipore) jusqu'à un volume inférieur à 500 μ L, puis injectée sur une colonne Superdex 75 (10x300 mm) à 0,8 mL/min à température ambiante, préalablement équilibrée avec du tampon Tris 20 mM pH 8, NaCl 150 mM. Les fractions d'éluion ont été analysées par SDS-PAGE et sur gel en conditions non-dénaturantes. Les fractions les plus pures en termes de tailles (SDS-PAGE) et les plus homogènes en termes de formes (gels non dénaturants) ont été regroupées et conservées à 4°C. Un tiers de pastille d'inhibiteurs de protéase CompleteTM (ROCHE) a été ajouté à chaque échantillon. Les 3 échantillons ont ensuite été dialysés contre 250 fois leur volume de tampon Tris 10 mM pH 8, NaCl 100 mM à 4°C, puis concentrés avec des unités de concentration Amicon (Millipore). Les concentrations ont été déterminées à partir de l'absorbance à 280 nm, en utilisant les coefficient molaire $\epsilon_{280} = 39500 \text{ cm}^{-1} \cdot \text{M}^{-1}$ pour $\alpha\beta\gamma$, $\epsilon_{280} = 17\,500 \text{ cm}^{-1} \cdot \text{M}^{-1}$ pour $\beta\gamma$, $\epsilon_{280} = 13\,950 \text{ cm}^{-1} \cdot \text{M}^{-1}$ pour β , déterminés par une analyse des acides aminés (Jean-Pierre Andrieux, LBM/IBS).

Les 3 échantillons concentrés ont ensuite été centrifugés pendant 40 minutes à 11 000 g à 4°C, puis transférés dans des tubes stériles. Les dilutions ont été réalisées avec le tampon de la dialyse filtré sur membrane Millipore 0,2 μ m. Tous les échantillons ont été conservés sur glace.

III.3. Immuno-révélation

Occasionnellement, l'identité des protéines purifiées a été vérifiée par western-blot. Environ 1 µg de protéine a été déposé sur SDS-PAGE 15%, puis transféré sur une membrane de nitrocellulose. L'immunorévélation a été réalisée après saturation de la membrane avec du lait 5% dans du PBS-tween 0,2%, par des anticorps polyclonaux anti-DivIB de lapin (dilution 1/4000^e, 1h d'incubation), anti-DivIC de souris (dilution 1/2000^e, 1h d'incubation), anti-FtsL de lapin (dilution 1/4000^e, 1h d'incubation) puis par des anticorps secondaires anti-lapins et anti-souris, couplés à l'enzyme HRP (dilution 1/10 000^e, incubation 1h).

IV. Expériences

IV.1. RMN

Le spectre HSQC [¹⁵N ; ¹H] présenté au chapitre 1/II.1, a été enregistré avec un échantillon EC/KL préparé comme décrit au paragraphe III.1, par mélange de protéine EC et KL (¹⁵N) exprimées et purifiées séparément (environ 10% de D₂O ont ensuite été ajoutés au volume de l'échantillon), sur un spectromètre Varian Innova 600 MHz, avec une largeur spectrale en hydrogène de 15,0 ppm et de 26,6 ppm en azote, 512 points complexes pour la dimension de l'hydrogène, 128 points complexes pour l'azote. Le signal de l'eau a été supprimé par sélection de cohérence par les gradients (spectre GHSQC).

Le spectre HSQC [¹⁵N ; ¹H] présenté au chapitre 2/I.1, a été enregistré sur un échantillon IB (ou αβγ), préparé comme décrit au paragraphe III.2.1, puis concentré avec une unité de concentration Amicon (Millipore), jusqu'à 0,14 mM (environ 10% de D₂O ont ensuite été ajoutés au volume de l'échantillon), sur un spectromètre Varian Innova 600 MHz, au bout de 1 heure, avec une largeur spectrale en hydrogène de 16,7 ppm et de 36,2 ppm en azote, 512 points complexes pour la dimension de l'hydrogène, 64 points complexes pour l'azote. Le signal de l'eau a été supprimé par sélection de cohérence par les gradients (spectre GHSQC).

Le spectre HSQC [¹⁵N ; ¹H] présenté au chapitre 2/I.2, a été enregistré sur un échantillon β, préparé comme décrit au paragraphe III.2.2, puis concentré avec une unité de concentration Amicon (Millipore), jusqu'à 0,27 mM (environ 10% de D₂O ont ensuite été ajoutés au volume de l'échantillon), sur un spectromètre Varian Innova 600 MHz, au bout de 40 minutes, avec une largeur spectrale en hydrogène de 16,7 ppm et de 36,2 ppm en azote,

512 points complexes pour la dimension de l'hydrogène, 128 points complexes pour l'azote. Le signal de l'eau a été supprimé par sélection de cohérence par les gradients (spectre GHSQC).

Le spectre HSQC [^{15}N ; ^1H] présenté au chapitre 3/1.1, a été enregistré sur l'échantillon de $\alpha\beta\gamma$ (ou IB), utilisé pour enregistrer le spectre HSQC présenté au chapitre 2/I.1, de 400 μL à 140 μM , auquel a été ajouté 100 μL d'un échantillon de EC/KL à 1200 μM préparé comme décrit au paragraphe III.1, sur un spectromètre Varian Innova 600 MHz, au bout de 1 heure, avec une largeur spectrale en hydrogène de 16,7 ppm et de 36,2 ppm en azote, 512 points complexes pour la dimension de l'hydrogène, 64 points complexes pour l'azote. Le signal de l'eau a été supprimé par sélection de cohérence par les gradients (spectre GHSQC).

Les spectres HSQC [^{15}N ; ^1H] présentés au chapitre 3/1.2, ont été enregistrés sur un échantillon de β préparé par Cécile Giustini (LRMN/IBS), à 400 μM , seul puis auquel 12,5 mg de complexe EC/KL préparé comme décrit au paragraphe III.1 et lyophilisé ont été resolubilisé, sur un spectromètre Varian Innova 600 MHz. Le spectre de la protéine β seule a été obtenu au bout de 15 minutes, avec une largeur spectrale en hydrogène de 12,6 ppm et de 29,6 ppm en azote, 301 points complexes pour la dimension de l'hydrogène, 64 points complexes pour l'azote. Le spectre de la protéine β complexée a été obtenu au bout de 63 heures, avec une largeur spectrale en hydrogène de 16,7 ppm et de 36,2 ppm en azote, 512 points complexes pour la dimension de l'hydrogène, 100 points complexes pour l'azote. Le signal de l'eau a été supprimé par sélection de cohérence par les gradients (spectres GHSQC).

IV.2. Diffusion dynamique de lumière (DLS)

Les échantillons EC/KL utilisés pour les expériences de DLS ont été préparés comme décrit au paragraphe III.1 par mélange de protéine EC et KL exprimées et purifiées séparément. Les mesures ont été réalisées, sur un appareil Dynapro-801TM, au Laboratoire de Cristallographie Moléculaire (IBS), dans lequel l'échantillon est illuminé par laser de 25 mW à 780 nm. La lumière diffusée est mesurée à un angle de 90°. Le rayon de Stokes a été déterminé à partir du coefficient de diffusion translationnel mesuré, par la formule :

Eq. 43 :

$$R_s = \frac{k_B T}{6\pi\eta D_T}$$

avec k_B , la constante de Boltzmann, T la température en Kelvin, mesurée par l'appareil, $\eta=1,02.10^{-3}$ N.s.m⁻¹, viscosité du tampon. Les erreurs sur chaque mesure de D_T sont moyennées, la moyenne des erreurs est donnée en pourcentage de la valeur du rayon de Stokes.

Le calcul du rayon de la sphère d'une protéine de 25 kDa a été fait avec la formule

Eq. 44 :

$$R_0 = \left(\frac{3M\bar{v}}{4\pi N_A} \right)^{\frac{1}{3}}$$

avec $M = 25$ kDa, $\bar{v} = 0,735.10^{-6}$ m³.g⁻¹ et N_A , la constante d'Avogadro.

IV.3. Chromatographie par exclusion de taille

L'expérience de chromatographie par exclusion de taille de EC/KL évoquée au chapitre 1/II.3, a été réalisée avec un échantillon de EC/KL, préparé comme décrit au paragraphe III.1 par mélange de protéine EC et KL exprimées et purifiées séparément. L'échantillon de 500 μ L a été injecté sur 24 mL de résine S200 (10x300 mm). Le volume d'élution était de 14,4 mL. La calibration de la colonne utilisée avait été réalisée le jour même. La droite résultant de la calibration était : $\ln(M_{app}) = -0,45v_{el} + 10,40$ où M_{app} est la masse apparente et v_{el} est le volume d'élution, avec un facteur R^2 de 0,99.

Les expériences de chromatographie par exclusion de taille de β , $\beta\gamma$ et $\alpha\beta\gamma$ (ou IB), présentées au chapitre 2/II.2, ont été réalisées avec des échantillons purifiés par chromatographie d'affinité (avec clivage de la protéine GST fusionnée) et par chromatographie échangeuse d'ions. La chromatographie par exclusion de taille des protéines β (2 échantillons de 500 μ L de 22 et 9 μ M) et $\alpha\beta\gamma$ (1 échantillon de 500 μ L de 24 μ M) a été réalisée sur 24 mL de résine S200 (10x300 mm). Les volumes d'élution de β étaient de 15,80 mL et 15,73 mL, celui de $\alpha\beta\gamma$ de 14,25 mL. La calibration de la colonne utilisée avait été réalisée la même semaine. La droite résultant de la calibration était : $\ln(M_{app}) = -0,51v_{el} + 11,25$ avec un facteur R^2 de 0,99. La chromatographie par exclusion de taille de $\beta\gamma$ (1 échantillon de 2 mL de 322 μ M) a été effectuée sur 120 mL de résine S200 (16x600 mm). Le volume d'élution de $\beta\gamma$ était de 88 mL. La calibration de la colonne avait été

faite 6 mois auparavant. La droite résultant de la calibration était $\ln(M_{app}) = -0,0766v_{el} + 11,191$ avec un facteur R^2 de 0,99.

IV.4. Ultracentrifugation analytique

Les expériences d'ultracentrifugation analytique ont toutes été réalisées au Laboratoire de Biophysique Moléculaire (IBS) avec Christine Ebel (LBM/IBS).

Les expériences d'ultracentrifugation analytique présentées au chapitre 1/II.3 ont été réalisées sur des échantillons de EC/KL préparés comme décrit au paragraphe III.1 par mélange de protéine EC et KL exprimées et purifiées séparément. La sédimentation à 130 000 g, 20°C a été suivie par l'absorbance à 280 nm, une mesure toutes les 6 minutes. L'analyse a été faite avec le logiciel Sedfit (Schuck, 2000), par ajustement de simili-courbes avec un coefficient de confiance de 0,7. La distribution en fonction du coefficient de sédimentation a été réalisée sur une gamme de S de 0,2 S à 20 S, avec un volume spécifique de 0,735 cm³/g, une densité du tampon à 1,005 et une viscosité à 0,0102 poise, estimés avec le logiciel Sednterp .

Les expériences d'ultracentrifugation analytique présentées au chapitre 1/V.1 ont été réalisées sur des échantillons de EC/KL préparés comme décrit au paragraphe III à partir de protéine EC et KL co-exprimées pour les échantillons totalement hydrogénés, et produit séparément pour les échantillons partiellement deutérés. La sédimentation à 130 000 g, 4°C a été suivie par interférence et par l'absorbance à 280 nm, une mesure toutes les 4 minutes. L'analyse a été faite avec le logiciel Sedfit, par ajustement de simili-courbes avec un coefficient de confiance de 0,7. La distribution en fonction du coefficient de sédimentation a été réalisée sur une gamme de S de 0,2 S à 20 S, avec un volume spécifique de 0,735 cm³/g, une densité du tampon à 1,0068 et une viscosité à 0,016 poise pour le tampon 100 % H₂O et une densité du tampon à 1,0511 et une viscosité à 0,016 poise pour le tampon 41% D₂O/ 59% H₂O, estimés avec le logiciel Sednterp.

L'expérience d'ultracentrifugation analytique présentée au chapitre 2/ II.2 a été réalisée sur un échantillon de $\alpha\beta\gamma$, ou IB, purifié par chromatographie d'affinité (avec clivage de la protéine GST fusionnée) de 460 μ M, à 130 000 g à 20°C. Les profils de sédimentation ont été suivis par mesure d'interférence, une mesure toutes les 6 minutes. L'analyse a été faite avec le logiciel Sedfit, par ajustement de simili-courbes avec un coefficient de confiance de 0,7. La distribution en fonction du coefficient de sédimentation et du rayon de stokes a été

réalisée sur une gamme de S de 0,2 S à 15 S , avec un rapport de coefficient de friction autorisé entre 1 et 2. Le volume spécifique a été estimé à 0,735 cm³/g, la densité du tampon à 1,005 et sa viscosité à 0,0102 poise avec le logiciel Sednterp.

IV.5. SANS et SAXS

IV.5.1. Enregistrement des courbes de diffusion

Les mesures de diffusion aux petits angles de neutrons ont été réalisées à l'Institut Laue Langevin, sur l'instrument D22, avec l'aide de Phil Callow (ILL) et Frank Gabel (LBM/IBS). La distance entre l'échantillon et le détecteur était de 2 et 8 m. La longueur d'onde des neutrons était de 6 Å. L'efficacité du détecteur a été calibrée par la diffusion d'eau déionisée, le bruit de fond électronique a été mesuré grâce à un échantillon de Boron (B₄C). Les échantillons ont été placés dans des cuves de quartz de 1 mm d'épaisseur, dans un porte-échantillon thermostaté à 4°C. Les clichés circulaires de diffusion ont ensuite été moyennés à partir du centre du faisceau, corrigés par la diffusion due au tampon, et celle de la cellule de quartz et par le bruit électronique (Gosh *et al.*, 2006).

Les mesures de diffusion aux petits angles de rayons X ont été réalisées à l'ESRF, sur la ligne ID02, avec l'aide de Anuj Shukla (ESRF), Stéphanie Finet (UPMC), Theyencheri Narayanan (ESRF) et Frank Gabel (LBM/IBS). La distance entre l'échantillon et le détecteur était de 2 m. La longueur d'onde des rayons X était de 0,1 nm et les données ont été collectées par un détecteur CCD. L'angle de diffusion et le centre du faisceau ont été calibrés par un échantillon de polyméthyl méthacrylate (PMMA) et un échantillon d'argent. Les échantillons ont été placés dans un capillaire de quartz de 2 mm environ de diamètre, et de 10 µm d'épaisseur de paroi environ, thermostaté à 15°C. Les dommages dus à la radiation ont été estimés par 10 expositions successives de 0,1 secondes d'un échantillon de chaque protéine. Le temps de mesure a été ensuite fixé à 1 secondes pour les échantillons de EC/KL et 0,1 secondes pour les échantillons de αβγ (ou IB), βγ et β. Nous avons mesuré les échantillons, du moins concentré au plus concentré, pour chaque protéine, en rinçant le capillaire avec 50 mL d'eau déionisée entre chaque échantillon. Environ 10 mesures ont été effectuées par échantillon, chaque mesure étant faite sur une fraction d'échantillon non encore exposée, en poussant celui-ci dans le capillaire entre chaque exposition. Les clichés circulaires de diffusion ont ensuite été moyennés à partir du centre du faisceau, normalisés par l'intensité du faisceau transmis et corrigés par la diffusion due au tampon.

IV.5.2. Extraction des valeurs d'intensité à l'origine et du rayon de giration

Les valeurs d'intensité à l'origine et du rayon de giration de chaque échantillon ont été extraites des courbes de diffusion avec le logiciel PRIMUS (Konarev *et al.*, 2003). Le logiciel permet l'ajustement d'une courbe de régression selon l'approximation de Guinier (cf. Introduction/B.II.2.1, Equation 14 et 15) à une représentation du logarithme de l'intensité en fonction du carré de Q. Le rayon de giration et l'intensité à l'origine sont extraits de la droite.

IV.5.3. Calibration de l'échelle d'intensité

Pour la calibration absolue de diffusion aux petits angles de neutrons, le calcul du facteur de proportionnalité f_1 a été effectué selon la formule

$$Eq. 45 : \quad f_1 = f \left(\frac{4\pi T_s}{1-T} \right) N_A t \left(\frac{\sum b - \rho_s V}{M} \right)^2$$

(Jacrot et Zaccai, 1981), où $f=0.8$ est un facteur de correction tabulé en fonction de la longueur d'onde utilisé ; $T_s=0,5$ et $T=0,5$ les facteurs de transmission du tampon et des échantillons ; $t=1$ mm l'épaisseur de la cuve ; et $\sum b - \rho_s V = 2,97 \cdot 10^{-14} \text{ cm.g}^{-1} \cdot \text{mol}$, le contraste.

Pour la calibration relative de diffusion aux petits angles de rayons X, des échantillons de lysozyme ont été préparés par pesée de 10 mg de protéine et dilution dans 1 mL du tampon filtré de dialyse des protéines. Les mesures des clichés de diffusion de l'échantillon se sont déroulées comme décrit au paragraphe précédent. Le facteur de proportionnalité a été calculé à partir de l'intensité à l'origine selon l'équation reformulée :

$$Eq. 46 : \quad f_2 = \frac{I_{0\text{exp}}}{M \cdot c}$$

Concentration c	Masse théorique M	$I(0)_{\text{exp}}$	Facteur f_2
10 g/l	14,3 kDa	$(11,8 \pm 0,1) \cdot 10^{-3}$	$(83 \pm 1) \cdot 10^{-6} \text{ mol.L.g}^{-2}$

Tableau 15 : Calcul du facteur de proportionnalité f_2 , entre l'intensité diffusée à l'origine et le produit de la masse et de la concentration pour les expériences de SAXS. Le calcul de ce facteur a été fait grâce à un échantillon de lysozyme de 10 mg/ml.

IV.5.4. Fonction de distribution de paires

L'estimation des fonction de distribution de paires a été réalisée avec le logiciel GNOM (Semenyuk et Svergun, 1991) à partir des courbes de diffusion des échantillons les plus concentrés dans la gamme de Q de 0,018 à 0,22 Å, pour les échantillons de EC/KL, de 0,012 à 0,25 pour $\beta\gamma$, de 0,013 à 0,19 pour $\beta\gamma$ /EC/KL.

IV.5.5. Calcul des simili-courbes de diffusion aux petits angles

Le calcul des simili-courbes de diffusion du domaine β d'*E. coli* et de celui de *G. stearothermophilus*, ont été calculés avec le logiciel CRY SOL (Svergun *et al.*, 1995) avec les paramètres par défaut du programme, à partir des structures déposées dans la Protein Data Bank, sous les références 2vh1 et 1yr1 respectivement. Pour le domaine β d'*E. coli*, les résidus 126 à 260 ont été considérés.

Les simili-courbes de diffusion des complexes construits de $\beta\gamma$ /EC/KL ont été calculées avec le logiciel CRYSON (Svergun *et al.*, 1998).

IV.5.6. Modélisation *ab initio*

Pour construire les enveloppes des molécules en trois dimensions, le programme de détermination *ab initio* DAMMIN (Svergun, 1999) a été utilisé, avec les paramètres par défaut du programme. Pour la protéine $\beta\gamma$, 6 modèles ont été construits à partir de la fonction de distribution de paires estimée avec GNOM et superposés ensuite pour vérifier la convergence du modèle. Les modèles ont été obtenus sans imposer de symétrie, à partir d'une sphère de diamètre équivalent à la distance maximale estimée avec GNOM. Pour le complexe $\beta\gamma$ /EC/KL, 5 modèles ont été construits à partir de la fonction de distribution de paires estimée avec GNOM et d'une sphère de rayon 160 Å. Puis 4 autres modèles ont été construits à partir d'un cylindre de 160 Å et de 40 Å de diamètre.

IV.5.7. Modélisation avec des corps rigides (RBM)

L'ajustement des corps rigides de manière automatique a été réalisé avec le logiciel MASSHA (Konarev *et al.*, 2001), avec le modèle de structure de la protéine $\beta\gamma$ et la structure de coiled-coil de la cortexilline (code pdb : 1d7m). L'ajustement a été effectué en partant de positions aléatoires et éloignées.

IV.6. Protéolyse limitée

L'expérience de protéolyse limitée présentée au chapitre 2/I.1, a été réalisée sur un échantillon de IB (ou $\alpha\beta\gamma$), purifié comme décrit au paragraphe III.2.4. Les fractions d'élution de la chromatographie échangeuse d'ions ont été mélangées, dialysées contre 200 fois leur volume de Tris 20 mM pH 8, NaCl 150 mM, puis concentrées avec une unité de concentration Amicon (Millipore) jusqu'à une concentration de 0,41 mM. Quatre mélanges réactionnels de $\alpha\beta\gamma$ (ou IB), ont été préparés en ajoutant 0,25 μ g de trypsine (Sigma T4549) à 25 μ g environ de protéine $\alpha\beta\gamma$. Un cinquième mélange contrôle a été préparé sans trypsine. Les 4 mélanges avec trypsine ont été incubés à 37°C, et les réactions ont été stoppées, pour le premier après 5 minutes, le second 25, le troisième 50 et le dernier après 75 minutes, par ajout de tampon de chargement pour SDS-PAGE et chauffage à 100°C. Après l'incubation, la moitié du volume des échantillons a été analysée par SDS-PAGE 15%.

IV.7. Biacore

IV.7.1.Acquisition des sensorgrammes

Les expériences d'analyse d'interaction de biomolécules par résonance plasmonique de surface ont été réalisées sur un instrument Biacore 3000 au Laboratoire d'Enzymologie Moléculaire avec l'assistance de Nicole Thiellens. Les échantillons de (his6)EC/(his6)KL, (his6)EC*/(his6)KL et (his6)EC/(his6)KL* ont été purifiés comme décrit au paragraphe III.1.3, les échantillons de β , $\beta\gamma$ et $\alpha\beta\gamma$ ont été purifiés comme décrit au paragraphe III.2. Tous les échantillons ont ensuite été dialysés contre 200 fois leur volume de Hepes 10 mM pH 7,5, NaCl 150 mM, EDTA 50 μ M, P20 0,005%. Le bio senseur utilisé était une surface NTA (Biacore), activée par 10 μ L de NiCl_2 500 μ M, et régénérée par 30 μ L de EDTA 350 mM, Hepes 10 mM pH 7,5, NaCl 150 mM, P20 0,005 %. Les expériences ont été réalisées à un débit de 20 μ L/min à température ambiante. Les échantillons « ligands », i.e. (his6)EC/(his6)KL, (his6)EC*/(his6)KL et (his6)EC/(his6)KL*, ont été dilués jusqu'à une concentration de 100 nM, les échantillons « analytes », i.e. β , $\beta\gamma$ et $\alpha\beta\gamma$, ont été dilués jusqu'à des concentrations de 0,005 à 1,5 mM, avec le tampon de dialyse filtré. Ce tampon a été utilisé comme tampon de course pour toutes les expériences. Pour chaque courbe, le biosenseur a été activé, puis les ligands ont été injectés. Le volume injecté a été adapté afin que les niveaux de fixation soient comparables pour chaque ligand. Ensuite, les analytes ont été injectés et la dissociation a été suivie pendant 5 minutes après la fin de l'injection. Enfin, le biosenseur a été

régénéré. Pour chaque expérience, 3 courbes ont été mesurées : la courbe d'interaction de l'analyte (β , $\beta\gamma$ ou $\alpha\beta\gamma$) sur le ligand (EC/KL, EC*/KL ou EC/KL*), la courbe d'interaction du tampon de course sur le ligand et la courbe d'interaction de l'analyte avec le capteur, en absence de ligand.

IV.7.2. Traitement des sensorgrammes

Les sensorgrammes ont ensuite été analysés avec le logiciel BIAevaluation 3.0. Pour chaque expérience, la courbe d'interaction du tampon sur ligand et la courbe d'interaction de l'analyte sur tampon ont été soustraites à la courbe d'interaction de l'analyte sur ligand. L'ajustement a ensuite été effectué sur une série de courbes d'interaction de $\beta\gamma$ à différentes concentrations sur EC/KL corrigées par les deux courbes de contrôle comme indiqué ci-dessus, avec un modèle d'interaction de Langmuir. Plusieurs ajustements ont été effectués (10) pour vérifier la reproductibilité d'obtention des valeurs de constante de dissociation et d'erreur. Les valeurs étaient systématiquement comprises entre 200 μM et 240 μM pour la constante de dissociation et entre 2 et 3 RU, pour l'erreur.

Discussion

I. Le complexe contraint EC/KL

Nous avons vu au chapitre 1 que le complexe contraint EC/KL est bien un hétéro complexe, formé d'une quantité équimolaire de protéine EC et KL. Nous avons vu que ce complexe était en équilibre entre un complexe tétramérique $(EC/KL)_2$ et dimérique $(EC/KL)_1$. Une constante de dissociation pour l'association de $(EC/KL)_1$ en $(EC/KL)_2$, a pu être estimée de l'ordre de 25 μ M, grâce aux données de diffusion aux petits angles de rayons X. Cette valeur est à utiliser avec précaution : elle a été obtenue avec peu de points et l'erreur sur ces points est assez importante du fait de la technique utilisée. Cependant, la diffusion aux petits angles aux rayons X est la seule technique qui a permis d'estimer cette valeur de dissociation, car c'est la seule où des paramètres du complexe $(EC/KL)_1$ monomérique, ont pu être mesurés. Pour toutes les autres techniques utilisées, les conditions où le complexe $(EC/KL)_1$ était seul en solution étaient en dessous des seuils de détection des instruments.

Il n'a pas été possible d'obtenir des informations structurales exploitables sur le complexe $(EC/KL)_1$. Nous n'avons pas trouvé de conditions d'observation du complexe $(EC/KL)_1$ compatibles avec aucune des méthodes utilisées.

Cependant, des informations sur le complexe $(EC/KL)_2$ ont pu être apportées : ce complexe n'est pas compact et sa forme semble plus proche de celle d'un cylindre que de celle d'une sphère. La distance maximale du complexe est comprise entre 170 et 200 Å environ. Il est possible que la difficulté d'estimer cette valeur soit due à la présence de segments flexibles. Les 2/3 (66 %, 143 résidus) de la séquence du complexe EC/KL sont prédits pour former une structure « coiled-coil ». La séquence N-terminale est composée des étiquettes de purification (10 % ; 21 résidus). Les séquences en C-terminal (24 % ; 51 résidus), sont sans doute impliquées dans l'interaction avec la partie extracellulaire de DivIB (cf. Chapitre 3/II.2). Il est possible que ces séquences N- et C-terminales ne soient pas structurées et flexibles dans le complexe $(EC/KL)_2$, et induisent une variabilité de la distance maximale du complexe (Figure 58).

Dans le complexe tétramérique $(EC/KL)_2$, la répartition des protéines EC et KL sont semblables : leurs profils de distributions de distances déterminés par SANS, sont très similaires. De plus, la forme globale de ces protéines au sein du complexe est proche de la

forme du complexe lui-même. Ces deux points sont cohérents avec une structure « coiled-coil », où les deux protéines sont juxtaposées longitudinalement, et s'enroulent légèrement l'une autour de l'autre. Leurs formes individuelles seraient aussi très similaires à celle du « coiled-coil » entier.

Plusieurs modèles d'auto-association du complexe (EC/KL)₁ en (EC/KL)₂ sont possibles : bout-à-bout, parallèles, bout-à-bout décalés... Les configurations proposées postulent une association faisant intervenir la complémentarité de charges des peptides e5 et k5.

Une structure « coiled-coil » de 200 acides aminés mesure 150 Å de long (cortexilline, code PDB : 1d7m). Un « coiled-coil » de 143 acides aminés (taille de la région de EC/KL prédite pour former un coiled-coil) mesurerait donc 107 Å environ. Dans le cas d'une association bout-à-bout (les deux « coiled-coil » à la suite l'un de l'autre), la structure du complexe (EC/KL)₂ mesurerait, au minimum, 214 Å (2 x 107), sans compter la contribution des séquences N- et C-terminales (Figure 58). Cela ne semble pas compatible avec la gamme des valeurs de la distance maximale du complexe.

Dans le cas d'une association en parallèle, la structure mesurerait au moins 107 Å (les deux « coiled-coil » l'un à côté de l'autre). En considérant les séquences en N- et C-terminales (72 résidus) comme totalement étirées, la chaîne peptidique de ces séquences mesurerait, au maximum, 139 Å (3,6 Å / résidu). Ainsi, dans le cas d'une structure du complexe (EC/KL)₂ en parallèle, la structure mesurerait au maximum 246 Å (107 Å, du coiled-coil + 139 Å, des chaînes non structurées totalement étirées) (Figure 58). Le cas d'une chaîne peptidique totalement étirée étant extrême, cette dimension est surestimée et donc compatible avec la gamme de valeurs de la distance maximale du complexe.

Il est aussi possible que la structure de (EC/KL)₂ soit une association de deux complexes (EC/KL)₁ bout-à-bout, mais décalés l'un par rapport à l'autre. Les deux « coiled-coil » seraient accolés, sur une partie de leur longueur (Figure 58).

Il est important d'estimer la pertinence des investigations sur le complexe tétramérique (EC/KL)₂. En effet, cette association de deux complexes (EC/KL)₁ reflète-t-elle une réalité physiologique ? Il est possible d'en douter, car cette auto-association est probablement due aux peptides artificiels e5 et k5. La présence de tétramère dans un mélange de e5 et k5, en plus du dimère attendu, avait été remarquée dès l'élaboration de ces peptides artificiels. La constante de dissociation de ce tétramère en dimères avait été estimée à 215 µM (Chao *et al.*, 1996). Cette valeur est supérieure à la constante de dissociation du complexe (EC/KL)₂, qui est de 25 µM environ. Les domaines extracellulaires des protéines DivIC et FtsL ont donc

peut-être un effet stabilisant sur la formation de $(EC/KL)_2$. Cet éventuel effet stabilisant traduit-il une réalité biologique ?

Dans tous les cas, il sera important de vérifier *in vivo*, l'existence de la formation d'un éventuel tétramère entre les protéines DivIC et FtsL, avant de pousser plus avant l'utilisation de ce modèle, qui a pu introduire un artefact du fait de l'utilisation de peptides artificiels.

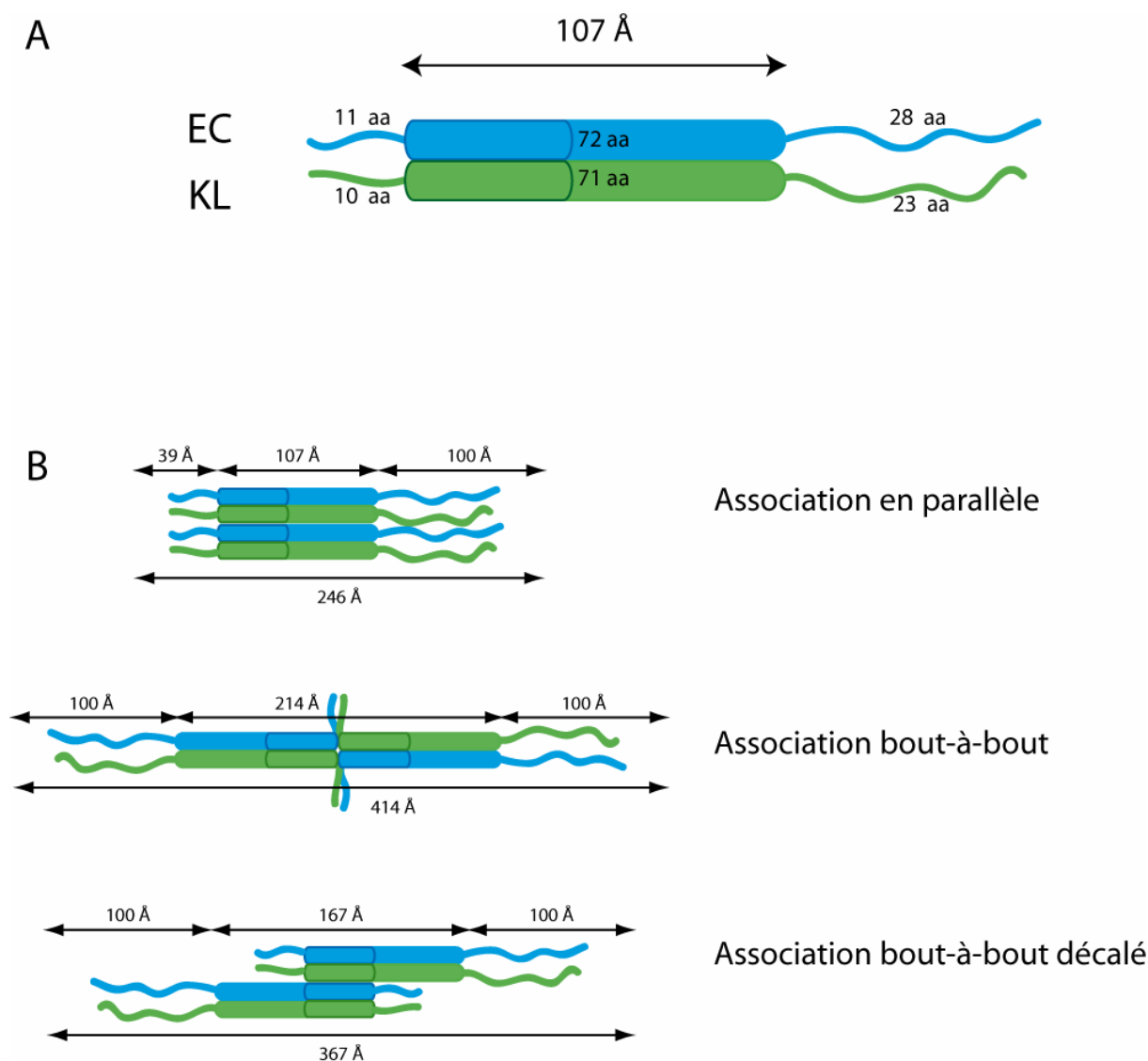


Figure 58 : **Modèles d'association des protéines EC et KL dans le complexe tétramérique $(EC/KL)_2$.** Les protéines EC (en bleu) et KL (en vert) sont prédites pour former un coiled-coil sur 72 acides aminés de leur séquence primaire, dont 32 acides aminés d'un peptide artificiel (délimité par un trait foncé dans le tube représentant le « coiled-coil »). Le nombre d'acides aminés des différentes régions de ces protéines est indiqué, ainsi que la longueur du « coiled-coil » (A). Différents modèles d'association du tétramère sont proposés, les longueurs maximales des différentes régions non-structurées sont indiquées au-dessus de ces régions (B).

II. La partie extracellulaire de DivIB

Dans le chapitre 2, nous avons ajouté des informations structurales sur la protéine DivIB de *S. pneumoniae* aux données déjà publiées portant sur les protéines de *E. coli* et *G. stearothermophilus*. Chez ces deux espèces, la protéine DivIB est constituée de trois domaines α , β et γ . Rappelons que les limites de domaines de la protéine DivIB ont été définies par l'examen de la structure cristallographique de *E. coli*, et par protéolyse limitée chez *G. stearothermophilus*. Chez *S. pneumoniae*, nous avons montré que cette protéine DivIB est également constituée de 3 domaines, par protéolyse limitée.

Un alignement des séquences primaires de la protéine DivIB de ces espèces de bactéries permet d'estimer la conservation des limites de ces domaines. L'alignement présenté a été réalisé en considérant la superposition des éléments de structures secondaires des structures de DivIB de *G. stearothermophilus* (code PDB : 1yr1) et d'*E. coli* (code PDB : 2vh1) (alignées avec le logiciel PYMOL).

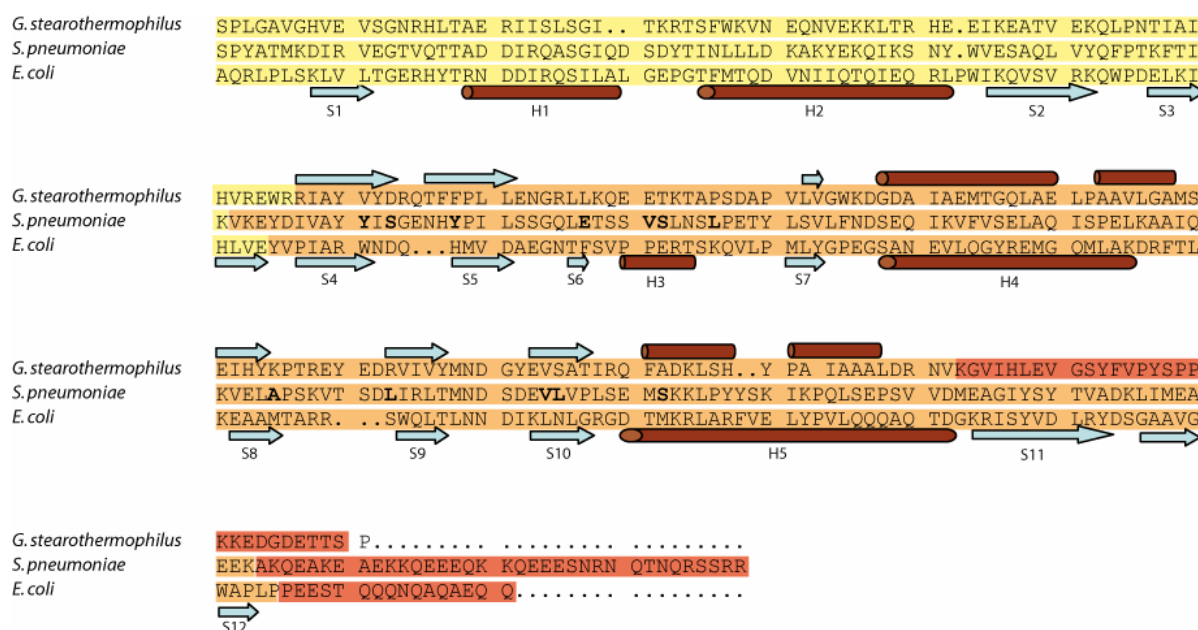


Figure 59 : Alignement manuel des séquences primaires et des éléments de structures secondaires des parties extracellulaires de DivIB de *G. stearothermophilus*, d'*E. coli*, et de *S. pneumoniae*. Les domaines α sont encadrés en jaune, les domaines β en orange, les domaines γ , en rouge. Les éléments de structures secondaires sont symbolisés par des flèches bleues pour les feuillets bêta et des tubes marrons pour les hélices alpha. La structure de la protéine d'*E. coli* est ainsi rapportée au dessus de sa séquence primaire, celle de *G. stearothermophilus*, en dessous de sa séquence primaire. Pour réaliser cet alignement, les éléments de structure secondaires ont été pris en compte. Les acides aminés identifiés par RMN comme étant impliqués dans l'interaction avec le complexe EC/KL sont indiqués en gras.

Les limites du domaine α , et la longueur de ce domaine sont très similaires chez ces trois espèces (en jaune sur la figure 59). La limite N-terminale correspond à la fin du segment transmembranaire prédit pour chacune de ces trois espèces. Les légères différences de la limite C-terminales sont sans doute dues d'une part, à la spécificité de la trypsine utilisée pour la protéolyse, et à un choix arbitraire dans le cas d'*E. coli*, reposant sur l'estimation de la fin du domaine structural. On peut remarquer que l'estimation de cette limite C-terminale du domaine α chez *E. coli* inclut le glutamate particulièrement conservé chez toutes les espèces de bactéries (Figure 59).

La similarité de séquences entre les espèces (Sanchez-Pulido *et al.*, 2003) indique que la structure de ce domaine doit être conservée, ce qui ne semble pas le cas. En effet, ce domaine est bien structuré pour la protéine recombinante d'*E. coli*, et celle d'*Y. enterocolitica* (van den Ent *et al.*, 2008), en revanche, il ne l'est pas dans la protéine recombinante de *G. stearothermophilus*, ni dans notre cas, dans la protéine recombinante de *S. pneumoniae*. Néanmoins, il est raisonnable de penser que ce domaine est structuré *in vivo*, mais que la stabilité de cette structure est faible, et dépende fortement de son environnement physiologique. Dans le cas de *S. pneumoniae* et *G. stearothermophilus*, la stabilité conformationnelle du domaine α est donc insuffisante pour permettre son repliement autonome dans une protéine recombinante soluble.

Il est probable que c'est la non structuration du domaine α qui ait induit l'auto-association de la partie extracellulaire de DivIB remarquée dans tous les échantillons de cette protéine. Cependant, certaines expériences d'interaction par double hybride chez la levure (Karimova *et al.*, 2005) et chez *E. coli* (Di Lallo *et al.*, 2003) suggèrent également une auto-association de la protéine DivIB. Cette question de l'interaction de DivIB avec elle-même est donc encore ouverte.

Dans cet alignement (Figure 59), les limites du domaine β de *S. pneumoniae* sont similaires à celui d'*E. coli* (en orange sur la figure), et donc plus larges que celles du domaine β de *G. stearothermophilus*. En effet, le domaine β de *G. stearothermophilus* est plus court de 21 acides aminés que les deux autres protéines.

Chez *G. stearothermophilus*, la délimitation des domaines a été faite par protéolyse limitée, et en considérant un alignement prenant en compte la conservation des éléments structuraux (Figure 59), on peut voir que la coupure enzymatique de la protéine de *G. stearothermophilus*, a lieu dans une région qui correspond dans la structure d'*E. coli* à une

boucle relativement exposée au solvant (entre l'hélice H5 et le brin S11). Il est donc possible que la protéase ait clivé la chaîne peptidique de la protéine de *G. stearothermophilus*, bien que les acides aminés en C-ter de cette coupure aient été structurés. Ainsi, les acides aminés C-terminaux de la protéine de *G. stearothermophilus* n'apparaîtraient pas dans la structure du domaine β bien que faisant partie structurellement de ce domaine. Il est également probable que cette séquence manquante de *G. stearothermophilus* se structure de manière comparable aux brins S11 et S12 de la séquence d'*E. coli*.

Nous avons également comparé la forme de la structure du domaine β de *S. pneumoniae* aux structures des domaines β de *G. stearothermophilus* et d'*E. coli*. Comme nous l'avons vu dans l'introduction (Introduction/A.V.4.3.2), les structures du domaine β de *G. stearothermophilus* et celle d'*E. coli* sont différentes. Nous avons montré ici (Chapitre 2/II.4.2) que la structure du domaine β de *S. pneumoniae* est très proche de celle du domaine β d'*E. coli*. La différence de forme avec la structure de *G. stearothermophilus* peut en partie s'expliquer par la différence de longueur du domaine. Le fait que la forme de la structure de *G. stearothermophilus* soit différente de la forme de la structure du même domaine chez deux autres espèces de bactérie confirme l'idée que les limites du domaine β de *G. stearothermophilus* définies par protéolyse limitée, n'englobent pas l'intégralité de ce domaine. Nous suggérons donc, sur la base de l'alignement présenté en figure 59, que la limite structurale du domaine β de *G. stearothermophilus* corresponde à celle de *S. pneumoniae* et d'*E. coli*, c'est-à-dire aux acides aminés terminaux de la séquence de DivIB chez *G. stearothermophilus*.

Nous avons également vu en introduction (cf Introduction/ V.4.3), que les deux brins bêta N-terminaux des structures du domaine β d'*E. coli* et de *G. stearothermophilus* ne sont pas identiquement positionnés dans les structure. En ce qui concerne ce point, il est difficile de conclure avec nos données. En effet, la forme du domaine β de *S. pneumoniae* est certes très proche de celle d'*E. coli*, néanmoins, la comparaison des courbes de diffusion n'est pas suffisamment précise pour distinguer ce niveau de détails. Cependant, il est possible que la structuration de ces deux brins bêta N-terminaux dans la structure de *G. stearothermophilus* soit influencée par l'absence du domaine α . La structure cristallographique d'*E. coli* est donc peut-être plus fiable.

En ce qui concerne le domaine γ , les résultats de ce travail indiquent que ce domaine n'est pas structuré chez *S. pneumoniae*, ce qui est d'ailleurs suggéré par sa composition (16 résidus chargés, 13 résidus polaires et 3 alanines) (en rouge dans la figure 59). Chez *E. coli*, ce

domaine n'est composé que de 16 acides aminés qui ne semblent pas structurés non plus, car ils sont absents de la structure cristallographique (van den Ent *et al.*, 2008). Chez *G. stearothermophilus*, si l'on considère une limite structurale du domaine β , définie par comparaison avec les autres espèces, le domaine γ ne comporterait que 5 acides aminés environ. Ainsi, il est possible que la fonction de ce domaine γ , qui est constitué d'une séquence d'acides aminés non-structurés, non-conservés et de longueur variable, soit peu importante. Elle est d'ailleurs inconnue à ce jour (Le Gouellec *et al.*, 2008).

III. Etude de l'interaction entre le complexe EC/KL et la partie extracellulaire de DivIB

Dans le chapitre 3, nous avons étudié l'interaction entre le complexe contraint EC/KL et la partie extracellulaire de DivIB. Nous avons vu que le complexe formé par ces trois protéines est anisotrope. Un modèle à basse résolution du complexe formé par le complexe EC/KL et la protéine $\beta\gamma$, confirme cette anisotropie. Dans ce modèle on peut distinguer une zone de masse excentrée, qui correspond vraisemblablement à la protéine $\beta\gamma$. La protéine $\beta\gamma$ se localiserait donc à une extrémité du complexe EC/KL, comme cela a été corroboré par l'invalidation de modèles où la protéine $\beta\gamma$ était localisée au centre d'une structure en « coiled-coil ».

III.1. Du côté du complexe EC/KL

L'interaction entre le complexe EC/KL et la protéine $\beta\gamma$ se fait via l'extrémité du complexe EC/KL ou plus exactement via les domaines C-terminaux de ces protéines, non concernés par la prédiction en « coiled-coil ». On a pu voir que l'un et l'autre de ces domaines (celui de DivIC et celui de FtsL) sont essentiels à la formation du complexe avec la partie extracellulaire de DivIB. Il avait été montré que la présence des parties extracellulaires des deux protéines DivIC et FtsL était nécessaire pour la formation d'un complexe avec la partie extracellulaire de DivIB (Noirclerc-Savoye *et al.*, 2005). Nous avons montré ici, par des expériences d'interaction observées par résonance plasmonique de surface, que c'est la

présence des régions C-terminales, non prédites en « coiled-coil », des deux protéines DivIC et FtsL qui est nécessaires.

Il est intéressant de remarquer que la mutation *div-355* dans le génome de la souche de *B. subtilis* thermosensible présentant des défauts de division à la température non-permissive, concerne ce domaine C-terminal de la protéine DivIC. Cette mutation décale le cadre de lecture du gène *divIC* à partir du codon de la méthionine 110 (cf. Introduction /A.V.2 ; Figure 8), et modifie la séquence des 15 derniers acides aminés et en rajoute 7 (Levin et Losick, 1994). Le domaine C-terminal correct de la protéine DivIC est donc essentiel à la division chez *B. subtilis*. Dans un alignement de séquence de DivIC, on remarque que deux acides aminés (une arginine et un glutamate, R101 et E111 chez *S. pneumoniae*) sont conservés dans les séquences primaires de ce domaine C-terminal (alignement présenté en figure 8, Introduction/V.2). De même pour la protéine FtsL, trois acides aminés (une arginine, une alanine et une isoleucine ; R84, I88 et A89 chez *S. pneumoniae*) sont conservés dans cette région C-terminale (alignement présenté en figure 7, Introduction/V.1.2). Cette conservation suggère que ces acides aminés sont importants pour la fonction de ces deux protéines DivIC et FtsL. Cette importance des régions C-terminales des protéines DivIC et FtsL pour leur fonction, a été remarqué *in vivo*, chez *B. subtilis* (Daniel *et al.*, 2006).

Nos résultats de diffusion aux petits angles indiquent que l'interaction du complexe EC/KL, avec la partie extracellulaire de DivIB induisait la dissociation du tétramère (EC/KL)₂ en dimère (EC/KL)₁. Deux protéines $\beta\gamma$ ne peuvent interagir simultanément avec le complexe (EC/KL)₂. Cela est peut-être dû à des contraintes stériques. On peut alors penser que dans le tétramère (EC/KL)₂, les deux zones d'interaction constituée des domaines C-terminaux des protéines EC et KL sont proches. Cela indiquerait que le tétramère est constitué de deux complexes (EC/KL)₁ associés en parallèle, où les domaines C-terminaux sont proches (Figure 58). Le fait que seul (EC/KL)₁ s'associe avec $\beta\gamma$ est un argument contre une éventuelle réalité physiologique de (EC/KL)₂.

III.2. Du côté de la partie extracellulaire de DivIB

Du côté de la partie extracellulaire de DivIB, les expériences de RMN indiquent que l'interaction se fait via le domaine β . Les autres domaines, α et γ , non structurés dans notre étude, ne se structurent pas au contact du complexe EC/KL, et ne semblent donc pas importants pour l'interaction avec le complexe EC/KL (cf. Chapitre 3/ I.1). Néanmoins, ce

dernier point a été observé sur un spectre de RMN, sur lequel il est difficile de déterminer précisément le nombre de pics dans la zone des pics non-structurés. Il n'est pas exclu que quelques acides aminés de ces domaines soient concernés par l'interaction avec le complexe EC/KL, mais certainement pas la plus grande partie de ces domaines qui contiennent 70 (α) et 35 (γ) acides aminés. Le modèle à basse résolution, obtenu à partir des données de diffusion aux petits angles de rayons X, indique aussi une interaction faisant intervenir le domaine β . Il est clair dans le modèle basse résolution du complexe $\beta\gamma$ /EC/KL, que l'interaction ne se fait pas uniquement via le domaine γ (cf. Chapitre 3/III.3.1 ; Figure 54).

Certains acides aminés du domaine β de DivIB de *S. pneumoniae* ont pu être identifiés comme impliqués dans l'interaction avec le complexe EC/KL. Ces acides aminés sont les tyrosines 229 et 236, les valines 249 et 312, les leucines 254, 301 et 313 (résidus non polaires), les sérines 231 et 250, et le glutamate 245 (résidus polaires). Rapportés sur une structure connue de ce domaine, ces acides aminés indiquent un épitope d'interaction, sur la face interne de la structure, au bout N-terminal (cf. Chapitre 3/I.2 ; Figure 45). Ces résultats sont cohérents avec l'orientation de la protéine $\beta\gamma$ dans le modèle à basse résolution du complexe $\beta\gamma$ /EC/KL, proposé dans la figure 55 (cf. Chapitre 3/III.3.1).

Quatre acides aminés de la protéine DivIB ont été identifiés par mutagenèse *in vivo* comme étant impliqués dans l'interaction avec les protéines DivIC et FtsL. Ces quatre acides aminés sont situés sur le brin S12, et l'hélice H5 de la structure d'*E. coli*, sur l'extrémité C-terminale de la structure (van den Ent *et al.*, 2008). L'attribution des ces éléments structuraux par RMN sur la protéine de *S. pneumoniae* n'étant pas terminée, ces acides aminés n'ont pas encore pu être identifiés comme interagissant avec le complexe contraint EC/KL. On peut toutefois remarquer que les deux acides aminés identifiés par mutagenèse *in vivo* sur le brin S12 (A252 et L259 dans la séquence d'*E. coli*) sont situés dans la face interne de la structure, et donc dans le prolongement de l'épitope déterminé par RMN.

Ces acides quatre aminés identifiés par mutagenèse chez *E. coli* sont, dans la séquence primaire et dans la structure, très proche de la queue γ . Chez *S. pneumoniae*, les 25 résidus C-terminaux du domaine β (correspondant aux deux brins S11 et S12 de la structure d'*E. coli*) sont importants pour la fonction de DivIB, probablement via l'interaction avec FtsL (Le Gouellec *et al.*, 2008). Cette région de la séquence (i.e. la région C-ter du domaine β), mais aussi le domaine γ , sont donc à considérer pour l'interaction avec FtsL et DivIC, malgré l'apparente non structuration du domaine γ .

III.3. A propos de l'interaction entre EC/KL et la partie extracellulaire de DivIB

Une constante de dissociation a pu être mesurée entre le complexe EC/KL et la protéine $\beta\gamma$. La valeur de cette constante, 0,22 μM , indique que ce complexe trimérique n'est pas extrêmement stable. Le complexe peut donc être vu comme une association qui peut être dissociée aisément. Cela est cohérent avec l'idée que, *in vivo*, ce complexe ne se forme que de manière transitoire (Noirclerc-Savoie *et al.*, 2005). Il est toutefois important de rester prudent quant à cette valeur qui a été mesurée sur un système modèle. Il est possible qu'elle ne reflète pas l'association des protéines natives, *in vivo*, notamment parce que les protéines natives sont membranaires. L'association entre ces trois protéines pourrait être stabilisée par les segments transmembranaires, ou contrainte par une concentration locale plus élevée (surfactive et non plus volumique), dans le cas où elles seraient côte à côte, dans la même membrane.

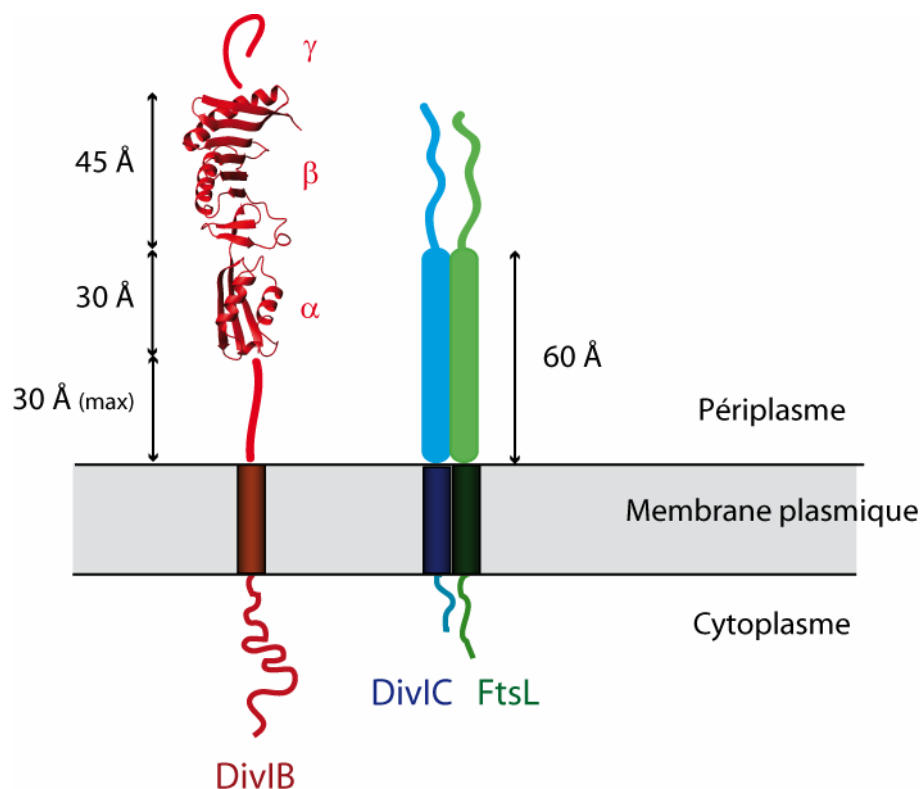


Figure 60 : Schéma de modèles structuraux des protéines DivIB, DivIC et FtsL, côte à côte dans une membrane. Les longueurs rapportées de la partie extracellulaire de DivIB (en rouge) ont été mesurées avec le logiciel PYMOL sur la structure d'*E. coli* (code PDB : 2vh1). Les 9 résidus proches de la membrane non présents dans la structure, et la queue γ , sont symbolisés par un trait rouge. La longueur de la structure en « coiled-coil » supposée formée par 40 acides aminés des parties extracellulaires des protéines DivIC (en bleu) et FtsL (en vert), a été calculée à partir de la structure de la cortexilline (code PDB 1d7m).

La distance à la membrane du domaine β est au moins la longueur du domaine α , c'est à dire de 30 Å environ. Neuf résidus de la séquence primaire de la partie extracellulaire de la protéine DivIB ne sont pas visibles dans la structure cristallographique de cette protéine d'*E. coli*. La longueur de la chaîne peptidique de ces 9 résidus serait au maximum de 30 Å (3,6 Å / résidu). La distance à la membrane du domaine β serait alors de 60 Å au maximum. Le domaine β est long de 45 Å environ.

La distance à la membrane des régions C-terminales des protéines DivIC et FtsL serait la longueur d'un « coiled-coil » de 40 acides aminés, c'est à dire de 60 Å. En effet, la prédiction en « coiled-coil » des protéines DivIC et FtsL concerne 40 acides aminés à partir des segments transmembranaires.

Ainsi, la structure en « coiled-coil » exposerait les régions C-terminales de FtsL et DivIC, 60 Å au dessus de la membrane. Dans cette situation, il est tout à fait possible que les domaines C-terminaux des protéines DivIC et FtsL interagissent avec la partie C-terminale du domaine β de DivIB, distant de 75 Å environ de la membrane (30 Å, pour le domaine α + 45 Å, pour le domaine β), si les 9 acides aminés proches de la membrane ne sont pas étirés.

Dans le cas d'une interaction des protéines DivIC et FtsL avec l'extrémité N-terminale du domaine β , les distances à la membrane sont aussi compatibles, mais plus difficilement. La zone d'interaction suggérée par les expériences de RMN est distante de la membrane de 70 Å au maximum, si les résidus proches de la membrane sont étirés. Cette valeur de distance comprend les 30 Å du domaine α , 10 Å de l'extrémité du domaine β à l'épitope d'interaction et 30 Å pour les 9 résidus de la séquence de DivIB entre le segment transmembranaire et le début du domaine α . Toutefois, les configurations ci-dessus considèrent les trois protéines orientées perpendiculairement à la membrane. Si les protéines peuvent être inclinées dans une certaine mesure par rapport au plan membranaire, une plus grande liberté conformationnelle d'assemblage est permise.

On peut également imaginer que les protéines DivIB, DivIC et FtsL ne soient pas côte à côte dans la membrane. Dans une situation de membrane plasmique invaginée, une zone de l'espace périplasmique présenterait deux faces membranaires en regard l'une de l'autre, où les domaines extracellulaires des protéines DivIB, DivIC, FtsL seraient en contact, sans que leur domaines transmembranaires soient proches. Cette situation a été suggérée par la difficulté à orienter la protéine $\beta\gamma$ dans le modèle à basse résolution du complexe trimérique $\beta\gamma$ /EC/KL (cf. Chapitre 3/III.3.1). En effet, la résolution du modèle, obtenu par modélisation *ab initio*, n'est pas suffisante pour orienter définitivement le modèle de la protéine $\beta\gamma$ par rapport au

complexe EC/KL (cf. figure 56). Si la protéine $\beta\gamma$ était orientée avec la queue γ dirigée vers la base du complexe EC/KL, la protéine DivIB serait orientée de manière anti-parallèle par rapport au complexe formé par les protéines DivIC et FtsL (Figure 61).

Cette situation est compatible avec les données de diffusion aux petits angles du complexe trimérique $\beta\gamma$ /EC/KL, mais elle imposerait une distance au septum entre les deux membranes relativement courte (200 Å au maximum), qui n'est peut-être pas réaliste (Matias et Beveridge, 2007; Zuber *et al.*, 2006).

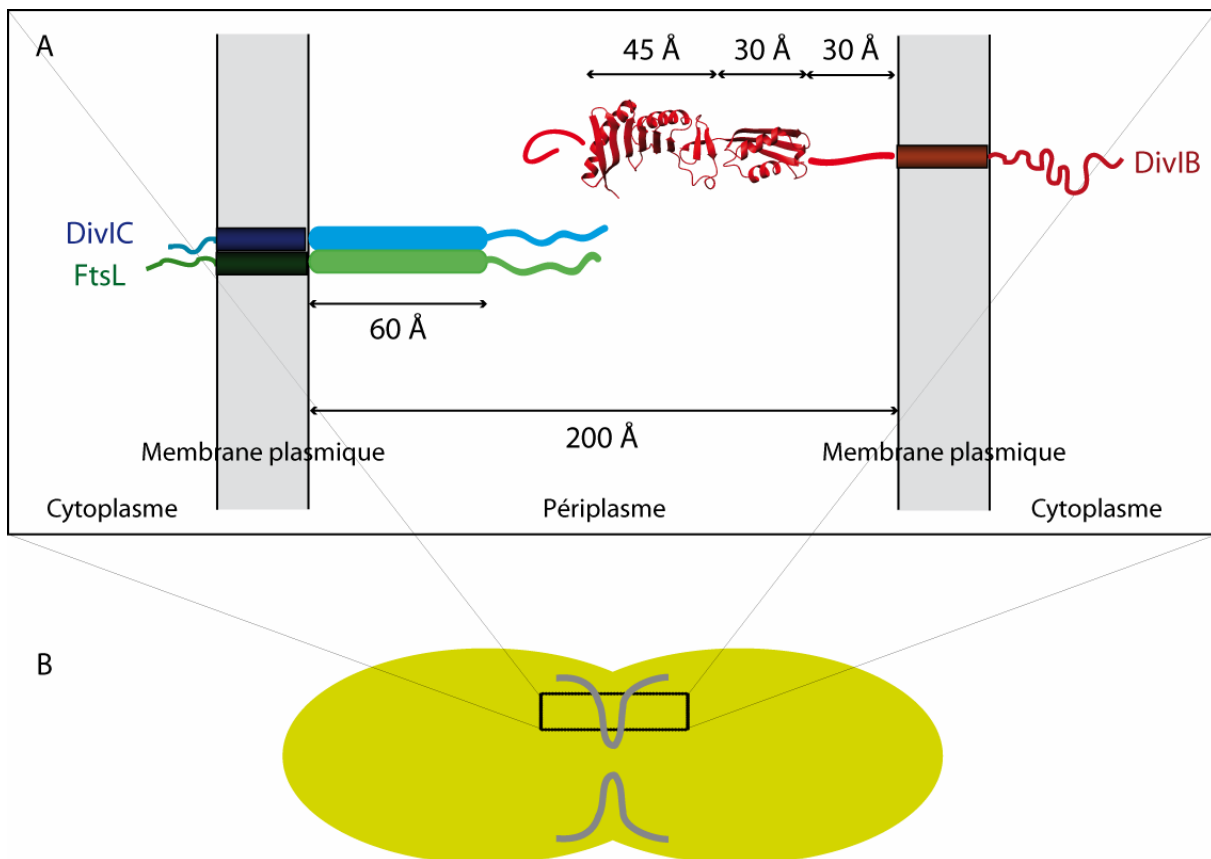


Figure 61 : Schéma d'un modèle d'association entre les protéines DivIB, DivIC et FtsL, face à face de part et d'autre de l'invagination, et localisation d'une telle situation sur le schéma d'une bactérie. Au site de division d'une bactérie, deux membranes se font faces, dans une zone du périplasm (B). Dans cette situation, un modèle d'association, face à face des protéines DivIB (en rouge), DivIC (en bleu) et FtsL (en vert) est proposé (A). Les longueurs rapportées de la partie extracellulaire de DivIB ont été mesurées avec le logiciel PYMOL sur la structure d'*E. coli* (code PDB : 2vh1). La longueur de la structure en « coiled-coil » supposée formée par 40 acides aminés des parties extracellulaires des protéines DivIC et FtsL, a été calculée à partir de la structure de la cortexilline (code PDB : 1d7m).

Conclusion et perspectives

Nos études structurales et biochimiques, qui ont fait appel à diverses méthodes biophysiques telles que la RMN, la diffusion aux petits angles, la résonance plasmonique de surface, nous ont permis de progresser dans notre compréhension du système constitué par les trois protéines DivIB, DivIC et FtsL.

Nous avons confirmé que la partie extracellulaire de la protéine DivIB de *S. pneumoniae* était constituée de trois domaines, α , β et γ . Notre modèle de structure du domaine β , obtenu par diffusion aux petits angles de rayons X est plus proche de la structure cristallographique du domaine β d'*E. coli* que du modèle obtenu par RMN du domaine β de *G. stearothermophilus*. Nos résultats ont ainsi montré que la structure de la protéine d'*E. coli* est vraisemblablement un meilleur modèle pour la partie extracellulaire de DivIB de *S. pneumoniae* que le modèle de la structure de *G. stearothermophilus*, bien que la similarité de séquence entre ces deux dernières espèces soit plus grande.

In vitro, la partie extracellulaire de DivIB interagit avec un complexe formé des parties extracellulaires des protéines DivIC et FtsL. Nos résultats indiquent que la partie extracellulaire de DivIB interagit via ce domaine central, β . Sur ce domaine β , un épitope d'interaction a été proposé par attribution de variations de déplacements chimiques obtenues par RMN. Cet épitope se situe sur la face interne de la structure d'*E. coli*, dans la région N-terminale de ce domaine β . L'interaction de plusieurs variants tronqués de la protéine DivIB sur un complexe formé des parties extracellulaires de DivIC et FtsL, a été testée par analyse d'interaction de biomolécules par SPR. Il est souhaitable de tester d'autres variants, notamment des mutants ponctuels au niveau de l'interface présumée d'interaction, et de comparer leur affinité par la même technique, ou encore d'observer *in vivo* l'impact de telles mutations. Une comparaison des constantes d'association/dissociation entre ces variants permettrait d'estimer l'importance relative des régions de la protéine DivIB dans l'interaction avec les protéines DivIC et FtsL.

D'autre part, les expériences qui nous ont permis d'identifier ces acides aminés du domaine β ont été réalisées sur le domaine β seul. Il serait intéressant de répéter ces expériences avec une protéine constituée des deux domaines β et γ de la partie extracellulaire de DivIB, dans le cas où ce domaine γ ait une influence sur une éventuelle interaction des

acides aminés C-terminaux du domaine β avec le complexe formé des parties extracellulaires de DivIC et FtsL.

Les protéines DivIC et FtsL interagissent avec la protéine DivIB via leur région C-terminale. Un complexe formé des régions centrales de ces deux protéines seulement, n'interagit pas avec la partie extracellulaire de DivIB. La présence des deux régions C-terminales de FtsL et DivIC est essentielle pour permettre l'interaction avec la partie extracellulaire de DivIB. Des résidus conservés ont été identifiés dans la séquence primaire de ces régions C-terminales. Il serait intéressant par la suite, de tester l'importance de ces acides aminés. Des tests d'interaction similaires à ceux présenté au chapitre 3, (analyse d'interaction de biomolécules par SPR), pourraient être réalisés, avec des protéines dans lesquelles ces acides aminés auraient été mutés, afin d'en évaluer l'importance. De même, *in vivo*, il serait souhaitable de tester par mutagenèse l'importance des résidus conservés pour la division et la survie des bactéries.

Nous avons proposé une première modélisation de la structure du complexe ternaire, formé des domaines β et γ de DivIB et d'un complexe des parties extracellulaires de DivIC et FtsL. Nous avons pu montrer que ce complexe se forme à l'extrémité d'une structure allongée qui est très vraisemblablement le complexe formé des parties extracellulaires de DivIC et FtsL. L'investigation de la structure de ce complexe formé par les parties extracellulaires de DivIC et FtsL n'a pas pu être poursuivie par RMN, à cause de la taille et de la forme allongée du complexe. D'autre part, ce complexe contraint des parties extracellulaires de DivIC et FtsL forme un tétramère à haute concentration, qui n'est pas probablement pas fonctionnel, notamment parce qu'il n'interagit pas avec la partie extracellulaire de DivIB. Ce système artificiel semble donc peu approprié pour poursuivre plus avant une étude structurale des protéines DivIC et FtsL.

La topologie des protéines DivIC et FtsL et leur interaction avec un troisième partenaire à leur extrémité, rappellent une structure particulière appelée « leucine-zipper » (Ellenberger *et al.*, 1992). En effet, une région de ces protéines forme une structure en « coiled-coil » qui permet leur dimérisation, et les régions N-terminales (C-terminales dans le cas de DivIC et FtsL) prolongent le « coiled-coil », et interagissent avec un partenaire (l'ADN dans le cas des « leucine-zipper »). Il serait intéressant de résoudre la structure des brins C-terminaux des protéines DivIC et FtsL dans l'interaction avec la partie extracellulaire de DivIB. En utilisant les propriétés de contraste de la diffusion aux petits angles de neutrons et un complexe de protéines partiellement deutéré, on pourrait tenter de déterminer comment ces

brins C-terminaux de DivIC et FtsL interagissent avec la partie extracellulaire de DivIB. On pourrait tenter de distinguer l'arrangement dans l'interface de ces brins C-terminaux de DivIC et FtsL dans le complexe ternaire et peut-être confirmer l'analogie avec un motif en « leucine-zipper » inversé.

Dans le but d'obtenir un modèle de structure à plus haute résolution, l'utilisation d'un système composé de peptides correspondants à ces brins C-terminaux de DivIC et FtsL et du domaine β de DivIB pourrait permettre de résoudre les problèmes de résolution des spectres de RMN, liés aux dimensions du complexe. En effet, un complexe du domaine β avec les deux peptides, pontés chimiquement de manière à mimer le rôle de dimérisation du « coiled-coil », permettrait sans doute l'acquisition de spectres de qualité, et pourrait également être adapté à une étude par cristallographie de rayons X.

Il serait également très intéressant de poursuivre l'étude du système utilisé dans ce travail, par diffusion aux petits angles de rayons X, afin d'augmenter la résolution du modèle de structure proposé. Nous avons proposé un modèle de structure du complexe formé par les domaines β et γ de DivIB, et d'un complexe des parties extracellulaires de DivIC et FtsL. La résolution de ce modèle n'est pas suffisante pour orienter de manière non ambiguë les domaines β et γ de DivIB par rapport au complexe formé des parties extracellulaires de DivIC et FtsL. Cette ambiguïté laisse ouverte la possibilité que les protéines DivIB, DivIC et FtsL interagissent en étant face à face, dans des plans membranaires parallèles. Cette hypothèse nous a permis de proposer un modèle d'association original, où ces protéines auraient un rôle au moment de l'invagination de la membrane, lors de la division. Il est donc important de valider (ou d'invalidier) ce modèle qui pourrait renseigner de manière inattendue sur la place et le rôle de ces protéines dans la division bactérienne.

En conclusion, grâce à l'utilisation de ce système modèle, nous avons pu obtenir de nombreuses informations structurales. La poursuite d'une étude structurale sur ce système permettra d'obtenir des informations supplémentaires. Néanmoins, dans l'optique d'obtenir des informations fonctionnelles, l'utilisation d'un système plus proche des protéines natives serait sûrement plus adaptée. L'ensemble des informations structurales recueillies au cours de cette thèse, permet désormais d'aborder ce type de travail. Une étude sur les protéines membranaires solubilisées pourra certainement ouvrir sur une meilleure reconstitution des processus dans lesquelles ces protéines sont impliquées.

Bibliographie

- Adam, M., Damblon, C., Jamin, M., Zorzi, W., Dusart, V., Galleni, M., el Kharroubi, A., Piras, G., Spratt, B.G., Keck, W., *et al.* (1991). Acyltransferase activities of the high-molecular-mass essential penicillin-binding proteins. *The Biochemical journal* 279 (Pt 2), 601-604.
- Addinall, S.G., Bi, E., and Lutkenhaus, J. (1996). FtsZ ring formation in fts mutants. *J Bacteriol* 178, 3877-3884.
- Addinall, S.G., and Lutkenhaus, J. (1996). FtsA is localized to the septum in an FtsZ-dependent manner. *J Bacteriol* 178, 7167-7172.
- Aue, W.P., Bartholdi, E., and Ernst, R.R. (1976). 2-Dimensional Spectroscopy-Application to nuclear Magnetic resonance. *Journal of Chemical Physics* 64, 2229-2246
- Aussel, L., Barre, F.X., Aroyo, M., Stasiak, A., Stasiak, A.Z., and Sherratt, D. (2002). FtsK Is a DNA motor protein that activates chromosome dimer resolution by switching the catalytic state of the XerC and XerD recombinases. *Cell* 108, 195-205.
- Beall, B., and Lutkenhaus, J. (1989). Nucleotide sequence and insertional inactivation of a *Bacillus subtilis* gene that affects cell division, sporulation, and temperature sensitivity. *J Bacteriol* 171, 6821-6834.
- Beattie, J., Phillips, K., Shand, J.H., Szymanowska, M., Flint, D.J., and Allan, G.J. (2008). Molecular interactions in the insulin-like growth factor (IGF) axis: a surface plasmon resonance (SPR) based biosensor study. *Molecular and cellular biochemistry* 307, 221-236.
- Begg, K.J., Dewar, S.J., and Donachie, W.D. (1995). A new *Escherichia coli* cell division gene, ftsK. *J Bacteriol* 177, 6211-6222.
- Begg, K.J., Hatfull, G.F., and Donachie, W.D. (1980). Identification of new genes in a cell envelope-cell division gene cluster of *Escherichia coli*: cell division gene ftsQ. *J Bacteriol* 144, 435-437.
- BEH (2006). Bulletin épidémiologique hebdomadaire 1.
- Bennett, J.A., Aimino, R.M., and McCormick, J.R. (2007). *Streptomyces coelicolor* genes ftsL and divIC play a role in cell division but are dispensable for colony formation. *J Bacteriol* 189, 8982-8992.
- Bi, E.F., and Lutkenhaus, J. (1991). FtsZ ring structure associated with division in *Escherichia coli*. *Nature* 354, 161-164.
- Blumberg, P.M., and Strominger, J.L. (1972a). Five Penicillin-binding Components Occur in *BaciEZus subtilis* Membranes. *The Journal of biological chemistry* 247, 8107-8113.

- Blumberg, P.M., and Strominger, J.L. (1972b). Isolation by covalent affinity chromatography of the penicillin-binding components from membranes of *Bacillus subtilis*. *Proceedings of the National Academy of Sciences of the United States of America* 69, 3751-3755.
- Blumberg, P.M., and Strominger, J.L. (1974). Interaction of penicillin with the bacterial cell: penicillin-binding proteins and penicillin-sensitive enzymes. *Bacteriol Rev* 38, 291-335.
- Bork, P., Sander, C., and Valencia, A. (1992). An ATPase domain common to prokaryotic cell cycle proteins, sugar kinases, actin, and hsp70 heat shock proteins. *Proc Natl Acad Sci U S A* 89, 7290-7294.
- Boyle, D.S., Khattar, M.M., Addinall, S.G., Lutkenhaus, J., and Donachie, W.D. (1997). *ftsW* is an essential cell-division gene in *Escherichia coli*. *Molecular microbiology* 24, 1263-1273.
- Bramhill, D. (1997). Bacterial cell division. *Annual review of cell and developmental biology* 13, 395-424.
- Bramhill, D., and Thompson, C.M. (1994). GTP-dependent polymerization of *Escherichia coli* FtsZ protein to form tubules. *Proc Natl Acad Sci U S A* 91, 5813-5817.
- Bramkamp, M., Weston, L., Daniel, R.A., and Errington, J. (2006). Regulated intramembrane proteolysis of FtsL protein and the control of cell division in *Bacillus subtilis*. *Molecular microbiology* 62, 580-591.
- Brown, M.S., Ye, J., Rawson, R.B., and Goldstein, J.L. (2000). Regulated intramembrane proteolysis: a control mechanism conserved from bacteria to humans. *Cell* 100, 391-398.
- Buddelmeijer, N., and Beckwith, J. (2004). A complex of the *Escherichia coli* cell division proteins FtsL, FtsB and FtsQ forms independently of its localization to the septal region. *Molecular microbiology* 52, 1315-1327.
- Buddelmeijer, N., Judson, N., Boyd, D., Mekalanos, J.J., and Beckwith, J. (2002). YgbQ, a cell division protein in *Escherichia coli* and *Vibrio cholerae*, localizes in codependent fashion with FtsL to the division site. *Proc Natl Acad Sci U S A* 99, 6316-6321.
- Burton, B.M., Marquis, K.A., Sullivan, N.L., Rapoport, T.A., and Rudner, D.Z. (2007). The ATPase SpoIIIE transports DNA across fused septal membranes during sporulation in *Bacillus subtilis*. *Cell* 131, 1301-1312.
- Butler, J.C. (2004). Epidemiology of pneumococcal disease. In *The Pneumococcus* E.I. Tuomanen, T.J. Mitchell, D.A. Morrison, and B.G. Spratt, eds. (Washington D.C., ASM Press), pp. 148-168.
- Campbell, N.A. (1995). *Biologie* (Bruxelles, De Boeck).

- Chain, E., Florey, H., and Gardner, N. (1940). Penicillin as a chemotherapeutic agent. *The lancet* 226-228.
- Chao, H., Houston, M.E., Jr., Grothe, S., Kay, C.M., O'Connor-McCourt, M., Irvin, R.T., and Hodges, R.S. (1996). Kinetic study on the formation of a de novo designed heterodimeric coiled-coil: use of surface plasmon resonance to monitor the association and dissociation of polypeptide chains. *Biochemistry* 35, 12175-12185.
- Chen, J.C., Weiss, D.S., Ghigo, J.M., and Beckwith, J. (1999). Septal localization of FtsQ, an essential cell division protein in *Escherichia coli*. *J Bacteriol* 181, 521-530.
- Cooper, P.D. (1956). Site of action of radiopenicillin. *Bacteriological reviews* 20, 28-48.
- Crook, D.W., Brueggemann, A.B., Sleeman, K.L., and Peto, T.E.A. (2004). Pneumococcal carriage. In *The Pneumococcus*, E.I. Tuomanen, T.J. Mitchell, D.A. Morrison, and B.G. Spratt, eds. (Washington D.C., ASM Press), pp. 136-147.
- D'Ulisse, V., Fagioli, M., Ghelardini, P., and Paolozzi, L. (2007). Three functional subdomains of the *Escherichia coli* FtsQ protein are involved in its interaction with the other division proteins. *Microbiology (Reading, England)* 153, 124-138.
- Dai, K., and Lutkenhaus, J. (1991). *ftsZ* Is an Essential Cell Division Gene in *Escherichia coli*. *Journal of bacteriology* 173, 3500-3506.
- Daniel, R.A., and Errington, J. (2000). Intrinsic instability of the essential cell division protein FtsL of *Bacillus subtilis* and a role for DivIB protein in FtsL turnover. *Molecular microbiology* 36, 278-289.
- Daniel, R.A., Harry, E.J., Katis, V.L., Wake, R.G., and Errington, J. (1998). Characterization of the essential cell division gene *ftsL(yIID)* of *Bacillus subtilis* and its role in the assembly of the division apparatus. *Molecular microbiology* 29, 593-604.
- Daniel, R.A., Noirot-Gros, M.F., Noirot, P., and Errington, J. (2006). Multiple interactions between the transmembrane division proteins of *Bacillus subtilis* and the role of FtsL instability in divisome assembly. *J Bacteriol* 188, 7396-7404.
- Daniel, R.A., Williams, A.M., and Errington, J. (1996). A complex four-gene operon containing essential cell division gene *pbpB* in *Bacillus subtilis*. *J Bacteriol* 178, 2343-2350.
- Darwin, C. (1859). *L'Origine des espèces*.
- de Boer, P., Crossley, R., and Rothfield, L. (1992). The essential bacterial cell-division protein FtsZ is a GTPase. *Nature* 359, 254-256.
- den Blaauwen, T., de Pedro, M.A., Nguyen-Disteche, M., and Ayala, J.A. (2008). Morphogenesis of rod-shaped sacculi. *FEMS Microbiol Rev* 32, 321-344.

- Di Lallo, G., Fagioli, M., Barionovi, D., Ghelardini, P., and Paolozzi, L. (2003). Use of a two-hybrid assay to study the assembly of a complex multicomponent protein machinery: bacterial septosome differentiation. *Microbiology (Reading, England)* *149*, 3353-3359.
- Din, N., Quardokus, E.M., Sackett, M.J., and Brun, Y.V. (1998). Dominant C-terminal deletions of FtsZ that affect its ability to localize in *Caulobacter* and its interaction with FtsA. *Molecular microbiology* *27*, 1051-1063.
- Ehlert, K., and Holtje, J.V. (1996). Role of precursor translocation in coordination of murein and phospholipid synthesis in *Escherichia coli*. *J Bacteriol* *178*, 6766-6771.
- Ellenberger, T.E., Brandl, C.J., Struhl, K., and Harrison, S.C. (1992). The GCN4 basic region leucine zipper binds DNA as a dimer of uninterrupted alpha helices: crystal structure of the protein-DNA complex. *Cell* *71*, 1223-1237.
- Erickson, H.P. (1995). FtsZ, a prokaryotic homolog of tubulin? *Cell* *80*, 367-370.
- Errington, J., Daniel, R.A., and Scheffers, D.J. (2003). Cytokinesis in bacteria. *Microbiol Mol Biol Rev* *67*, 52-65, table of contents.
- Fleming, A. (1945). Penicillin: Nobel Lecture.
- Fleming, A. (2001). On the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of *B. influenzae*. 1929. *Bulletin of the World Health Organization* *79*, 780-790.
- Gerard, P., Vernet, T., and Zapun, A. (2002). Membrane topology of the *Streptococcus pneumoniae* FtsW division protein. *J Bacteriol* *184*, 1925-1931.
- Ghigo, J.M., and Beckwith, J. (2000). Cell division in *Escherichia coli*: role of FtsL domains in septal localization, function, and oligomerization. *J Bacteriol* *182*, 116-129.
- Ghigo, J.M., Weiss, D.S., Chen, J.C., Yarrow, J.C., and Beckwith, J. (1999). Localization of FtsL to the *Escherichia coli* septal ring. *Molecular microbiology* *31*, 725-737.
- Ghuysen, J.M., and Goffin, C. (1999). Lack of cell wall peptidoglycan versus penicillin sensitivity: new insights into the chlamydial anomaly. *Antimicrobial agents and chemotherapy* *43*, 2339-2344.
- Goranov, A.I., Katz, L., Breier, A.M., Burge, C.B., and Grossman, A.D. (2005). A transcriptional response to replication status mediated by the conserved bacterial replication protein DnaA. *Proc Natl Acad Sci U S A* *102*, 12932-12937.
- Gosh, R.E., Egelhaaf, S.U., and Rennie, A.R. (2006). A computing guide for small-angle scattering experiments. Institute Laue Langevin internal report ILL06GH05T.
- Guinier A., and Fournet, G. (1955). *Small Angle Scattering of X-rays* (John Wiley & Sons, New York London).

- Gutmann, L., and Varon, E. (2006). Rapport d'activité 2006 du Centre National de Référence des Pneumocoques.
- Guzman, L.M., Barondess, J.J., and Beckwith, J. (1992). FtsL, an essential cytoplasmic membrane protein involved in cell division in *Escherichia coli*. *J Bacteriol* *174*, 7716-7728.
- Hajduk, S.L. (1976). Demonstration of kinetoplast DNA in dyskinetoplastic strains of *Trypanosoma equiperdum*. *Science (New York, NY)* *191*, 858-859.
- Hara, H., Yasuda, S., Horiuchi, K., and Park, J.T. (1997). A promoter for the first nine genes of the *Escherichia coli* mra cluster of cell division and cell envelope biosynthesis genes, including ftsI and ftsW. *J Bacteriol* *179*, 5802-5811.
- Hirokawa, T., Boon-Chieng, S., and Mitaku, S. (1998). SOSUI: classification and secondary structure prediction system for membrane proteins. *Bioinformatics (Oxford, England)* *14*, 378-379.
- Hirota, Y., Ryter, A., and Jacob, F. (1968). Thermosensitive mutants of *E. coli* affected in the processes of DNA synthesis and cellular division. *Cold Spring Harbor symposia on quantitative biology* *33*, 677-693.
- Ishino, F., and Matsushashi, M. (1981). Peptidoglycan synthetic enzyme activities of highly purified penicillin-binding protein 3 in *Escherichia coli*: a septum-forming reaction sequence. *Biochemical and biophysical research communications* *101*, 905-911.
- Izaki, K., Matsushashi, M., and Strominger, J.L. (1968). Biosynthesis of the peptidoglycan of bacterial cell walls. 8. Peptidoglycan transpeptidase and D-alanine carboxypeptidase: penicillin-sensitive enzymatic reaction in strains of *Escherichia coli*. *J Biol Chem* *243*, 3180-3192.
- Jacrot, B. (1976). Study of biological structures by neutron-scattering from solution. *Reports on Progress in Physics* *39*, 911-953.
- Jacrot, B., and Zaccai, G. (1981). Determination of molecular-weight by neutron-scattering. *Biopolymers* *20*, 2413-2426.
- Karimova, G., Dautin, N., and Ladant, D. (2005). Interaction network among *Escherichia coli* membrane proteins involved in cell division as revealed by bacterial two-hybrid analysis. *J Bacteriol* *187*, 2233-2243.
- Katis, V.L., Harry, E.J., and Wake, R.G. (1997). The *Bacillus subtilis* division protein DivIC is a highly abundant membrane-bound protein that localizes to the division site. *Molecular microbiology* *26*, 1047-1055.
- Katis, V.L., Wake, R.G., and Harry, E.J. (2000). Septal localization of the membrane-bound division proteins of *Bacillus subtilis* DivIB and DivIC is codependent only at high temperatures and requires FtsZ. *J Bacteriol* *182*, 3607-3611.

- Kell, C.M., Sharma, U.K., Dowson, C.G., Town, C., Balganes, T.S., and Spratt, B.G. (1993). Deletion analysis of the essentiality of penicillin-binding proteins 1A, 2B and 2X of *Streptococcus pneumoniae*. *FEMS microbiology letters* 106, 171-175.
- Kendrew, J.C., Bodo, G., Dintzis, H.M., Parrish, R.G., Wyckoff, H., and Phillips, D.C. (1958). A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. *Nature* 181, 662-666.
- Konarev P. V., Petoukhov M. V., and Svergun, S. (2001). MASSHA - a graphics system for rigid-body modelling of macromolecular complexes against solution scattering data. *Journal of Applied Crystallography* 34, 527-532.
- Konarev P.V., V.V., Sokolova A.V., Koch M.H.J., Svergun D.I. (2003). PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *Journal of Applied Crystallography* 36, 1277-1282.
- Koradi, R., Billeter, M., and Wuthrich, K. (1996). MOLMOL: a program for display and analysis of macromolecular structures. *Journal of molecular graphics* 14, 51-55, 29-32.
- Lara, B., Rico, A.I., Petruzzelli, S., Santona, A., Dumas, J., Biton, J., Vicente, M., Mingorance, J., and Massidda, O. (2005). Cell division in cocci: localization and properties of the *Streptococcus pneumoniae* FtsA protein. *Molecular microbiology* 55, 699-711.
- Le Gouellec, A., Roux, L., Fadda, D., Massidda, O., Vernet, T., and Zapun, A. (2008). Roles of pneumococcal DivIB in cell division. *J Bacteriol* 190, 4501-4511.
- Lederberg, J. (1957). Mechanism of action of penicillin. *Journal of bacteriology* 73, 144.
- Levin, P.A., and Losick, R. (1994). Characterization of a cell division gene from *Bacillus subtilis* that is required for vegetative and sporulation septum formation. *J Bacteriol* 176, 1451-1459.
- Levitt, M.H. (2001). *Spin dynamics* (Wiley).
- Li, Z., Trimble, M.J., Brun, Y.V., and Jensen, G.J. (2007). The structure of FtsZ filaments in vivo suggests a force-generating role in cell division. *The EMBO journal* 26, 4694-4708.
- Ligon, B.L. (2004). Penicillin: its discovery and early development. *Seminars in pediatric infectious diseases* 15, 52-57.
- Liu, G., Draper, G.C., and Donachie, W.D. (1998). FtsK is a bifunctional protein involved in cell division and chromosome localization in *Escherichia coli*. *Molecular microbiology* 29, 893-903.
- Lowe, J., and Amos, L.A. (1998). Crystal structure of the bacterial cell-division protein FtsZ. *Nature* 391, 203-206.

- Lu, C., Reedy, M., and Erickson, H.P. (2000). Straight and curved conformations of FtsZ are regulated by GTP hydrolysis. *J Bacteriol* 182, 164-170.
- Lupas, A. (1996). Coiled coils: new structures and new functions. *Trends in biochemical sciences* 21, 375-382.
- Lutkenhaus, J. (1993). FtsZ ring in bacterial cytokinesis. *Molecular microbiology* 9, 403-409.
- Lutkenhaus JF, D.W. (1979). Identification of the ftsA Gene Product. *Journal of bacteriology* 137, 1088-1094.
- Ma, X., Ehrhardt, D.W., and Margolin, W. (1996). Colocalization of cell division proteins FtsZ and FtsA to cytoskeletal structures in living *Escherichia coli* cells by using green fluorescent protein. *Proc Natl Acad Sci U S A* 93, 12998-13003.
- Ma, X., and Margolin, W. (1999). Genetic and functional analyses of the conserved C-terminal core domain of *Escherichia coli* FtsZ. *J Bacteriol* 181, 7531-7544.
- Mancini, E.J., Kainov, D.E., Grimes, J.M., Tuma, R., Bamford, D.H., and Stuart, D.I. (2004). Atomic snapshots of an RNA packaging motor reveal conformational changes linking ATP hydrolysis to RNA translocation. *Cell* 118, 743-755.
- Margalit DN, Romberg L, Mets RB, Hebert AM, Mitchison TJ, Kirschner MW, and D., R. (2004). Targeting cell division Small-molecule inhibitors of FtsZ GTPase perturb cytokinetic ring assembly and induce bacterial lethality. *Proceedings of the National Academy of Sciences of the United States of America* 101, 11821-11826.
- Margolin, W. (2000). Themes and variations in prokaryotic cell division. *FEMS microbiology reviews* 24, 531-548.
- Massey, T.H., Mercogliano, C.P., Yates, J., Sherratt, D.J., and Lowe, J. (2006). Double-stranded DNA translocation: structure and mechanism of hexameric FtsK. *Molecular cell* 23, 457-469.
- Massidda, O., Anderluzzi, D., Friedli, L., and Feger, G. (1998). Unconventional organization of the division and cell wall gene cluster of *Streptococcus pneumoniae*. *Microbiology (Reading, England)* 144 (Pt 11), 3069-3078.
- Matias, V.R., and Beveridge, T.J. (2007). Cryo-electron microscopy of cell division in *Staphylococcus aureus* reveals a mid-zone between nascent cross walls. *Molecular microbiology* 64, 195-206.
- Mendelson, N., and Cole, R. (1972). Genetic Regulation of Cell Division Initiation in *Bacillus subtilis*. *Journal of bacteriology* 112, 994-1003.
- Mercer, K.L., and Weiss, D.S. (2002). The *Escherichia coli* cell division protein FtsW is required to recruit its cognate transpeptidase, FtsI (PBP3), to the division site. *J Bacteriol* 184, 904-912.

- Mitchell, P., and J. Moyle (1956). Osmotic function and structure in bacteria. *Symp Soc Gen Microbiol* 6:150.
- Mylonas E, Svergun D. (2007). Accuracy of molecular mass determination of proteins in solution by small-angle X-ray scattering. *Journal of Applied Crystallography* 40, S245-S249.
- Nishimura, Y., Takeda, Y., Nishimura, A., Suzuki, H., Inouye, M., and Hirota, Y. (1977). Synthetic ColE1 plasmids carrying genes for cell division in *Escherichia coli*. *Plasmid* 1, 67-77.
- Noirclerc-Savoye, M., Le Gouellec, A., Morlot, C., Dideberg, O., Vernet, T., and Zapun, A. (2005). In vitro reconstitution of a trimeric complex of DivIB, DivIC and FtsL, and their transient co-localization at the division site in *Streptococcus pneumoniae*. *Molecular microbiology* 55, 413-424.
- Pease, P.J., Levy, O., Cost, G.J., Gore, J., Ptacin, J.L., Sherratt, D., Bustamante, C., and Cozzarelli, N.R. (2005). Sequence-directed DNA translocation by purified FtsK. *Science (New York, NY)* 307, 586-590.
- Perutz, M.F. (1960). Structure of hemoglobin. *Brookhaven symposia in biology* 13, 165-183.
- Pichoff, S., and Lutkenhaus, J. (2005). Tethering the Z ring to the membrane through a conserved membrane targeting sequence in FtsA. *Molecular microbiology* 55, 1722-1734.
- Ricard, M., and Hirota, Y. (1973). Process of cellular division in *Escherichia coli*: physiological study on thermosensitive mutants defective in cell division. *J Bacteriol* 116, 314-322.
- Robson, S.A., and King, G.F. (2006). Domain architecture and structure of the bacterial cell division protein DivIB. *Proc Natl Acad Sci U S A* 103, 6700-6705.
- Robson, S.A., Michie, K.A., Mackay, J.P., Harry, E., and King, G.F. (2002). The *Bacillus subtilis* cell division proteins FtsL and DivIC are intrinsically unstable and do not interact with one another in the absence of other septasomal components. *Molecular microbiology* 44, 663-674.
- Rothschild, L.J., and Mancinelli, R.L. (2001). Life in extreme environments. *Nature* 409, 1092-1101.
- Saleh, O.A., Peral, C., Barre, F.X., and Allemand, J.F. (2004). Fast, DNA-sequence independent translocation by FtsK in a single-molecule experiment. *The EMBO journal* 23, 2430-2439.
- Sanchez-Pulido, L., Devos, D., Genevrois, S., Vicente, M., and Valencia, A. (2003). POTRA: a conserved domain in the FtsQ family and a class of beta-barrel outer membrane proteins. *Trends in biochemical sciences* 28, 523-526.

- Sanchez, M., Valencia, A., Ferrandiz, M.J., Sander, C., and Vicente, M. (1994). Correlation between the structure and biochemical activities of FtsA, an essential cell division protein of the actin family. *The EMBO journal* *13*, 4919-4925.
- Schuck, P. (2000). Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophysical journal* *78*, 1606-1619.
- Sears, V.F. (1992). *Neutron news*. *3*, 29-37.
- Semenyuk, A., and Svergun, D. (1991). GNOM - A program package for Small-Angle Scattering Data-Processing. *Journal of Applied Cristallography* *24*, 537-540.
- Serdyuck, I.N., Zaccai, N., and Zaccai, J. (2007). *Methods in Molecular Biophics* (Cambridge University Press).
- Siddiqui, R.A., Hoischen, C., Holst, O., Heinze, I., Schlott, B., Gumpert, J., Diekmann, S., Grosse, F., and Platzer, M. (2006). The analysis of cell division and cell wall synthesis genes reveals mutationally inactivated ftsQ and mraY in a protoplast-type L-form of *Escherichia coli*. *FEMS microbiology letters* *258*, 305-311.
- Sievers, J., and Errington, J. (2000). The *Bacillus subtilis* cell division protein FtsL localizes to sites of septation and interacts with DivIC. *Molecular microbiology* *36*, 846-855.
- Singleton, M.R., Sawaya, M.R., Ellenberger, T., and Wigley, D.B. (2000). Crystal structure of T7 gene 4 ring helicase indicates a mechanism for sequential hydrolysis of nucleotides. *Cell* *101*, 589-600.
- Spratt, B.G. (1975). Distinct penicillin binding proteins involved in the division, elongation, and shape of *Escherichia coli* K12. *Proc Natl Acad Sci U S A* *72*, 2999-3003.
- Spratt, B.G. (1977). Temperature-sensitive cell division mutants of *Escherichia coli* with thermolabile penicillin-binding proteins. *J Bacteriol* *131*, 293-305.
- Strick, T.R., and Quessada-Vial, A. (2006). FtsK: a groovy helicase. *Nature structural & molecular biology* *13*, 948-950.
- Subramanya, H.S., Bird, L.E., Brannigan, J.A., and Wigley, D.B. (1996). Crystal structure of a DExx box DNA helicase. *Nature* *384*, 379-383.
- Svergun, D. (1999). Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophysical journal* *77*, 2896-2896.
- Svergun D, B.C., Koch MHJ (1995). CRY SOL - A program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates. *Journal Of Applied Crystallography* *28*, 768-773.
- Svergun, D.I., Richard, S., Koch, M.H., Sayers, Z., Kuprin, S., and Zaccai, G. (1998). Protein hydration in solution: experimental observation by x-ray and neutron scattering. *Proc Natl Acad Sci U S A* *95*, 2267-2272.

- Tipper, D.J., and Strominger, J.L. (1965). Mechanism of action of penicillins: a proposal based on their structural similarity to acyl-D-alanyl-D-alanine. *Proceedings of the National Academy of Sciences of the United States of America* *54*, 1133-1141.
- Todar, K. (2007). *The Bacterial Flora of Humans*.
- Tugarinov, V., Kanelis, V., and Kay, L.E. (2006). Isotope labeling strategies for the study of high-molecular-weight proteins by solution NMR spectroscopy. *Nature protocols* *1*, 749-754.
- van den Ent, F., and Lowe, J. (2000). Crystal structure of the cell division protein FtsA from *Thermotoga maritima*. *The EMBO journal* *19*, 5300-5307.
- van den Ent, F., Vinkenvleugel, T.M., Ind, A., West, P., Veprintsev, D., Nanninga, N., den Blaauwen, T., and Lowe, J. (2008). Structural and mutational analysis of the cell division protein FtsQ. *Molecular microbiology* *68*, 110-123.
- van Heijenoort, J. (2007). Lipid intermediates in the biosynthesis of bacterial peptidoglycan. *Microbiol Mol Biol Rev* *71*, 620-635.
- Wang, L., and Lutkenhaus, J. (1998). FtsK is an essential cell division protein that is localized to the septum and induced as part of the SOS response. *Molecular microbiology* *29*, 731-740.
- Whitman, W., Coleman, D., and Wiebe, W. (1998). Prokaryotes: the unseen majority *Proc Natl Acad Sci U S A* *95* 6578 – 6583.
- Yan, K., Pearce, K.H., and Payne, D.J. (2000). A conserved residue at the extreme C-terminus of FtsZ is critical for the FtsA-FtsZ interaction in *Staphylococcus aureus*. *Biochemical and biophysical research communications* *270*, 387-392.
- Yu, X.C., Weihe, E.K., and Margolin, W. (1998). Role of the C terminus of FtsK in *Escherichia coli* chromosome segregation. *J Bacteriol* *180*, 6424-6428.
- Zapun, A., Vernet, T., and Pinho, M.G. (2008). The different shapes of cocci. *FEMS microbiology reviews* *32*, 345-360.
- Zuber, B., Haenni, M., Ribeiro, T., Minnig, K., Lopes, F., Moreillon, P., and Dubochet, J. (2006). Granular layer in the periplasmic space of gram-positive bacteria and fine structures of *Enterococcus gallinarum* and *Streptococcus gordonii* septa revealed by cryo-electron microscopy of vitreous sections. *J Bacteriol* *188*, 6652-6660.

Annexe : démonstration des formules utilisées

I. Démonstration de l'équation 35 p.79

On considère que toute la masse d'une molécule de masse molaire M et de volume spécifique \bar{v} est contenue dans une sphère de rayon R_0 .

$$Eq. 47 : \quad \frac{4\pi}{3} R_0^3 = \frac{M}{N_A} \bar{v}$$

Grâce à cette formule et à la formule 34, on exprime le coefficient de friction de la molécule idéalement sphérique, comme l'indique la formule 48.

$$Eq. 34 : \quad f = 6\pi\eta R$$

$$Eq. 48 : \quad f_0 = 6\pi\eta R_0 = 6\pi\eta \left(\frac{3M\bar{v}}{4\pi N_A} \right)^{\frac{1}{3}}$$

En utilisant la formule 32, on exprime le coefficient de sédimentation s en fonction des coefficients de friction réel et d'une molécule idéalement sphérique.

$$Eq. 32 : \quad s = \frac{M(1 - \bar{v}\rho)}{N_A f}$$

$$Eq. 49 : \quad M = \frac{s \cdot N_A f}{1 - \rho \bar{v}} \frac{f_0}{f} = \frac{s \cdot N_A}{1 - \rho \bar{v}} \frac{f}{f_0} 6\pi\eta R_0 = \frac{s \cdot N_A}{1 - \rho \bar{v}} \frac{f}{f_0} 6\pi\eta \left(\frac{3M\bar{v}}{4\pi N_A} \right)^{\frac{1}{3}}$$

La reformulation de l'équation 49 donne la formule 35.

$$Eq. 35 : \quad M_{app} = \left(\frac{s \cdot N_A \cdot 6\pi}{1 - \rho \bar{v}} \cdot \frac{f}{f_0} \right)^{\frac{3}{2}} \left(\frac{3\bar{v}}{4\pi N_A} \right)^{\frac{1}{2}}$$

II. Démonstration de la formule 39 p.131

La constante de dissociation du complexe ternaire s'exprime en fonction des concentrations de $\beta\gamma$, de $(EC/KL)_1$ et de $\beta\gamma/(EC/KL)_1$, comme l'indique la formule 50.

$$Eq. 50 : \quad K_{D1} = \frac{[\beta\gamma][(EC/KL)_1]}{[\beta\gamma/(EC/KL)_1]}$$

La constante de dissociation du tétramère $(EC/KL)_2$ s'exprime en fonction des concentrations du tétramère $(EC/KL)_2$ et du dimère $(EC/KL)_1$, comme l'indique la formule 51.

$$Eq. 51 : \quad K_{D2} = \frac{[(EC/KL)_1]^2}{[(EC/KL)_2]}$$

L'équation 42 de conservation peut s'exprimer en fonction des concentrations des espèces, comme l'indique les formules 52 et 53.

$$Eq. 42 : \quad [EC/KL]_{total} = [\beta\gamma]_{total} = C_0$$

$$Eq. 52 : \quad C_0 = [(EC/KL)_1] + 2[(EC/KL)_2] + [\beta\gamma/(EC/KL)_1]$$

$$Eq. 53 : \quad C_0 = [\beta\gamma] + [\beta\gamma/(EC/KL)_1]$$

Dans la formule 52, on injecte la formule 40, comme l'indique la formule 54.

$$Eq. 40 : \quad [(EC/KL)_2] = \frac{1}{K_{D2}} [(EC/KL)_1]^2$$

$$Eq. 54 : \quad C_0 = [(EC/KL)_1] + 2 \frac{[(EC/KL)_1]^2}{K_{D2}} + [\beta\gamma/(EC/KL)_1]$$

La reformulation de la formule 50 permet d'exprimer la concentration du complexe ternaire en fonction de celle de $\beta\gamma$ et de celle de $(EC/KL)_1$, comme l'indique la formule 55.

$$Eq. 55 : \quad [\beta\gamma/(EC/KL)_1] = \frac{[\beta\gamma][(EC/KL)_1]}{K_{D1}}$$

On injecte l'équation 55 dans l'équation 54 pour obtenir l'équation 56.

$$Eq. 56 : \quad C_0 = [(EC / KL)_1] + 2 \frac{[(EC / KL)_1]^2}{K_{D2}} + \frac{[\beta\gamma][(EC / KL)_1]}{K_{D1}}$$

La concentration de $\beta\gamma$ peut s'exprimer en fonction de celle de $(EC/KL)_1$, comme l'indique la formule 57, à partir de la formule 52, 53 et 40.

Eq. 57 :

$$[\beta\gamma] = C_0 - [\beta\gamma / (EC / KL)_1] = C_0 - (C_0 - [(EC / KL)_1] - 2[(EC / KL)_2]) = [(EC / KL)_1] + 2 \frac{[(EC / KL)_1]^2}{K_{D2}}$$

L'injection de la formule 57 dans la formule 56 permet d'obtenir l'équation 58.

$$Eq. 58 : C_0 = [(EC / KL)_1] + 2 \frac{[(EC / KL)_1]^2}{K_{D2}} + \frac{[(EC / KL)_1]}{K_{D1}} \left([(EC / KL)_1] + 2 \frac{[(EC / KL)_1]^2}{K_{D2}} \right)$$

La reformulation de l'équation 58 permet d'obtenir l'équation 39.

$$Eq. 39 : \quad \frac{2}{K_{D1}K_{D2}} [(EC / KL)_1]^3 + \left(\frac{1}{K_{D1}} + \frac{2}{K_{D2}} \right) [(EC / KL)_1]^2 + [(EC / KL)_1] = C_0$$

Résumé/Abstract

FtsL, DivIC et DivIB sont trois protéines membranaires impliquées dans la division bactérienne. Leur fonction n'est pas totalement comprise, mais semble mutuellement dépendante, notamment à travers la formation de complexes. Pour contribuer à la connaissance structurale des protéines de la division bactérienne et apporter des indices sur la fonction des trois protéines citées, une étude structurale a été menée sur un système modèle de protéines recombinantes solubles de *S. pneumoniae* : FtsL, DivIC et DivIB. La partie extracellulaire de DivIB, un complexe contraint des parties extracellulaires de FtsL et DivIC, et l'interaction entre ce complexe et la partie extracellulaire de DivIB ont été étudiés par plusieurs techniques biophysiques (RMN, SANS et SAXS, BIA par SPR). La partie extracellulaire de DivIB est composée de trois domaines dont le domaine central est structuralement proche de son orthologue chez *E. coli*, et interagit avec un complexe des parties extracellulaires de DivIC et FtsL *via* ce domaine central. Un épitope d'interaction sur ce domaine a été identifié. Les domaines C-terminaux de FtsL et DivIC sont essentiels à l'interaction avec la partie extracellulaire de DivIB. Un modèle à basse résolution du complexe de ces trois protéines, présente en effet le domaine central de la partie extracellulaire de DivIB à l'extrémité du complexe des parties extracellulaires de DivIC et FtsL. Différents modèles d'association dans la cellule, des protéines FtsL, DivIC et DivIB ont été évalués avec ces nouvelles données structurales.

FtsL, DivIC and DivIB are three membrane proteins involved in bacterial division. Their function remains uncertain, but is mutually dependant through the formation of complexes. To contribute to the structural knowledge of bacterial division proteins, and to bring insights on the function of the three cited proteins, we carried out a structural study on a model system of recombinant soluble proteins from *S. pneumoniae*: FtsL, DivIC and DivIB. We have studied by several biophysical techniques (NMR, SANS and SAXS, BIA by SPR) the extracellular part of DivIB, a constrained complex of the extracellular parts of FtsL and DivIC, and the interaction between this complex and the extracellular part of DivIB. The extracellular part of DivIB is composed of three domains. The central domain of DivIB is structurally close to its orthologue in *E. coli*. DivIB interacts with a complex of the extracellular parts of FtsL and DivIC, *via* this central domain. We have identified an epitope of interaction on this domain. The C-terminus domain of FtsL and DivIC are essentials for the interaction with the extracellular part of DivIB. Indeed, a low resolution model of the three proteins complex shows that the central domain of the extracellular part of DivIB is located at the tip of the complex of the extracellular parts of FtsL and DivIC. We have evaluated different models of the association of FtsL, DivIC and DivIB in the cell with these new structural data.

Mots-clés: *Streptococcus pneumoniae*, DivIB, DivIC, FtsL, division bactérienne, SANS, SAXS, RMN, Biacore

Key words: *Streptococcus pneumoniae*, DivIB, DivIC, FtsL, bacterial division, SANS, SAXS, RMN, Biacore